Appendix B – Evidence tables

Critical appraisal and findings tables

Views and experiences

Review question 1 – What are the views and experiences of children and young people, their caregivers and families, and adult survivors of child abuse in the UK on the process of recognising and assessing abuse and neglect, and on services providing early help for, or intervention following, abuse and neglect of children and young people?

Studies for this question are presented alongside the relevant review area (recognition, assessment, early help, response).

Review question 2 – What are the views and experiences of practitioners working in the UK on the process of recognising and assessing abuse and neglect, and on services providing early help for, or intervention following, abuse and neglect of children and young people?

Studies for this question are presented alongside the relevant review area (recognition, assessment, early help, response).

Recognition

Review question 3 – What emotional, behavioural and social (non-clinical) indicators relating to children and young people should alert practitioners to the possibility of abuse and neglect?

Review question 4 – What emotional, behavioural and social (non-clinical) indicators relating to caregivers and families should alert practitioners to the possibility of abuse and neglect?

Many of the papers reviewed contained information relevant to both of the above questions, and so are presented together.

Review questions 3 and 4 – Critical appraisal tables

1. **Allen B, Tussey C (2012) Can Projective Drawings Detect if a Child Experienced Sexual or Physical Abuse? A Systematic Review of the Controlled Research. Trauma, Violence, and Abuse: A Review Journal 13: 97–111**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: 'A comprehensive literature review of the controlled research to determine whether any graphic indicators (e.g., genitalia, omission of body parts) or predefined scoring system can reliability [sic] and validly discriminate abused from nonabused children' (p97). **Methodology:** Systematic review. **Appropriate and clearly focused question?** Yes. Use of drawings to determine the validity of sexual abuse allegations. **Adequate description of methodology?** Partly adequate. No report of quality assessment of data	**Inclusion of relevant individual studies?** Yes. Inclusion criteria were papers published prior to 2011, intervention group either sexual or physical abuse, but NOT combination of different types of abuse, used a control group (non-abused), clear evaluation criteria prior to data analysis. **Study quality assessed and reported?** Partly reported. Author stated that '… studies are evaluated in light of their methodological rigor, including interrater reliability, blinding of the raters to participant condition, the degree of match between the abused and control groups on extraneous factors (e.g., mental health status),	**Does the study's research question match the review question?** Yes. Use of drawings to determine the validity of sexual abuse allegations. **Has the study dealt appropriately with any ethical concerns?** Not reported. **Were service users involved in the study?** Not reported. **Is there a clear focus on the guideline topic?** Yes. Recognition of graphic indicators in sexual abuse allegations. **Is the study population the same as at least 1 of the groups**	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity rating:** - Inconsistent reporting of statistical data from original studies. No report of quality assessment of data extraction, discussion between reviewers.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
extraction, discussion between reviewers etc. **Rigorous literature search?** Yes. PsycINFO, MEDLINE and PILOTS databases were utilised. Keywords used were projective, sexual abuse, physical abuse, drawing, human figure drawing, Kinetic Family Drawing (KFD), House-Tree-Person, Draw-A-Person (DAP). Reference list of identified studies examined to identify additional relevant studies. Table of included studies presented	and statistical procedures employed' (p99). General comments on study quality given as '… quality of these studies varied widely and, accordingly, interpretations of the findings are often difficult' (p107). **Do conclusions match findings?** Partly. Limited methodological description of systematic review methodology.	**covered by the guideline?** Yes. Abused children. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Partly. Not applicable. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Recognition. **Does the study have a UK perspective?** No.	

2. **De Bellis MD, Hooper SR, Spratt EG et al. (2009) Neuropsychological findings in childhood neglect and their relationships to pediatric PTSD. Journal of the International Neuropsychological Society 15: 868–78**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: The objective was to examine impact of neglect on IQ, reading, maths, fine-motor skills, language, visual-spatial, memory/learning and attention/executive functions in 2 groups of non-sexually abused medically health neglect children, 1 with post-traumatic stress disorder and 1 without, and 1 nonmaltreated control group. Here we have extracted findings only in relation to impact on language.	**Measurements and outcomes clear?** Yes. Validated measures used for measuring all outcomes. For language measures (NEPSY and Peabody Picture Vocabulary test) reliability of scales not reported. **Measurements valid?** Yes. **Setting for data collection justified?** Yes.	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Yes. Study approved by local hospital institutional review board. 'Legal guardians gave informed consent and children assented prior to participation' (p80).	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + No justification given for age range of participants, and not made clear whether results apply

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: Cross-sectional study.[1] **Objectives of study clearly stated?** Yes. **Clearly specified and appropriate research design?** Yes. Observational comparative design, comparing children with neglect and post-traumatic stress disorder, neglect without post-traumatic stress disorder and non-neglected controls. **Subjects recruited in acceptable way?** Yes. Neglect groups were recruited '… through advertisements targeted at DSS agencies …' (p869). Control group participants 'recruited from the same surrounding community through Duke University Medical Centre Institutional Review Board (IRB)-approved advertisement at schools and pediatric clinics …' (p869). **Sample representative of defined population?** Partly. Unclear	**Are all important outcomes and results considered?** Yes. **Tables/graphs adequately labelled and understandable?** Yes. **Appropriate choice and use of statistical methods?** Yes. Multivariate analysis of four outcome measures relevant to language with follow-up pairwise comparisons. Bonferroni-corrected significance reported where criterion was met, but 0.05 used as criterion elsewhere. Would have been better to have used Bonferroni-corrected criterion throughout, to correct for multiple analyses. **In-depth description of the analysis process?** Partly. Unclear how variation in ages of participants in the three groups were taken in to account. **Are sufficient data presented to support the findings?** Yes. **Results discussed in relation to existing knowledge on the subject and study objectives?** Yes.	**Were service users involved in the study?** No. Service users were involved as participants, but not in design or interpretation of results. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children aged 7 to 13 who have experienced neglect with our without post-traumatic stress disorder, and a non-neglected control group. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to recognition. **Does the study have a UK perspective?** No. Study conducted in North Carolina, USA.	to this age range only. Whilst ages across the three groups were not statistically significantly different, there was variation in the mean age and age ranges across groups. This does not appear to be used as a covariate in the analysis. Unclear why only neglect and post-traumatic stress disorder group selected for within-group analysis, rather than all children who had experienced neglect.

[1] The term cross-sectional study is used here to denote an observational studies in which exposure and outcome are measured at the same time (that is, not longitudinal study). This is consistent with the terminology used in the NICE guideline development manual on study classification.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
why this is age group (7 to 13) selected rather than full range of children.	**Results can be generalised?** Partly. Results could be generalised to children of this age group. **Do conclusions match findings?** Yes.		

3. Eigsti I, Cicchetti D (2004) The impact of child maltreatment on expressive syntax at 60 months. Developmental Science 7: 88–102

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: The main aim of the study was to '… examine spontaneous language in a sample of maltreated children and well-matched comparison children, focusing specifically on the syntactic complexity of the children's utterances, to see whether the deficits observed in previous studies of maltreated toddlers extend to syntactic complexity in school-age children' (p92). A secondary goal of the study was determine if differences in maternal utterances were correlated with the syntactic development of children. **Methodology:** Cross-sectional study. **Objectives of study clearly stated?** Yes.	**Measurements and outcomes clear?** Yes. The study focused on language deficit and the methods and measures used to assess this are clear. **Measurements valid?** Partly. All assessments were conducted using pre-established measures, however data in relation to reliability and validity of these are not reported. These appear to be valid to the approach taken. Assessment was conducted by coding videotaped play interactions between the mother and her child. The authors report inter-rater reliability testing for the transcription of these recordings however this is not reported for the coding stage.	**Does the study's research question match the review question?** Yes. The main aim of the study was to '… examine spontaneous language in a sample of maltreated children and well-matched comparison children, focusing specifically on the syntactic complexity of the children's utterances, to see whether the deficits observed in previous studies of maltreated toddlers extend to syntactic complexity in school-age children' (p92). A secondary goal of the study was determine if differences in maternal utterances were correlated with the syntactic development of children. **Has the study dealt appropriately with any ethical concerns?** No. The authors do not report participant consent processes or note	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + Some lack of clarity in the paper regarding statistical analysis.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Clearly specified and appropriate research design? No. The authors simply state that the paper reports on a cross-sectional project which was part of a larger longitudinal/cross-sectional study (Harvard Child Maltreatment Project). **Subjects recruited in acceptable way?** Yes. The authors provide a reasonably adequate description of the recruitment process for both groups and these were acceptable. **Sample representative of defined population?** Unclear. The authors report that the '… sample was representative of all child protective cases in the greater Boston region at the time of the study' (Cicchetti & Manly 1990, p92). However it is unclear how they determined this.	**Setting for data collection justified?** Partly. The only detail provided in relation to setting is that assessments took place in a playroom – this seems likely to have been part of a research facility. **Are all important outcomes and results considered?** Yes. **Tables/graphs adequately labelled and understandable?** Yes. **Appropriate choice and use of statistical methods?** Partly. The authors report that, for child language data, MANCOVA was conducted with group (maltreated vs. comparison) as the independent variable and outcome measures as dependent variable. However, for some measures gender was also examined as an independent variable. It is not always clear whether this was achieved via 2-way MANCOVA with group and gender as independent variables, or sequential MANCOVAs. We have assumed throughout that 2-way MANCOVA was conducted, and highlighted where interaction effects were not reported. For some measures, 'effects' of SES are also reported – here we have	whether the study was approved by an ethics committee. **Were service users involved in the study?** No. Service users involved as participants only, no indication of involvement in design of study or interpretation of findings. **Is there a clear focus on the guideline topic?** Yes. The study focuses on language deficits in maltreated children. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Maltreated children and their mothers. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Recognition – indicators relating to children and young people. **Does the study have a UK perspective?** No. The study was conducted in the USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	assumed that a separate analysis was conducted, although this is not specified. **In-depth description of the analysis process?** Yes. **Are sufficient data presented to support the findings?** Yes. **Results discussed in relation to existing knowledge on the subject and study objectives?** Yes. **Results can be generalised?** Unclear. **Do conclusions match findings?** Yes.		

4. **Evans E, Hawton K, Rodham K (2005) Suicidal phenomena and abuse in adolescents: a review of epidemiological studies. Child Abuse and Neglect 29: 45–58**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. Investigating association between experiencing abuse and experiencing suicidal thoughts and behaviours. **Adequate description of methodology?** Partly adequate. Essentially narrative review, with in-	**Inclusion of relevant individual studies?** Yes. **Study quality assessed and reported?** No. The review does not critically appraise included studies. However, this was a common feature across the systematic reviews of observational data which we found, and there is recognition	**Does the study's research question match the review question?** Yes. Relates to Q3 on recognition. **Has the study dealt appropriately with any ethical concerns?** Not reported.	**Overall assessment of internal validity:** + The review does not critically appraise included studies. However, this was a common feature across the systematic reviews of observational data which we found, and there is recognition that methods

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
clusion of odds ratio data. The authors do not justify why they have not conducted a meta-analysis of studies, even though it appears that it would have been possible from the data they had available. **Rigorous literature search?** Yes. Seven bibliographic databases search, reference harvesting from identified papers and search of relevant websites. However, no hand searching of key journals.	that methods of critical appraisal for observational studies are less well developed. We have therefore included this study, and not 'marked down' the overall quality rating on this ground. **Do conclusions match findings?** Yes. Narrative discussion is in line with data presented, and draws out nuances and contradictions as appropriate.	**Were service users involved in the study?** No. Service users involved as participants in the included studies, but not included in systematic review itself. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Partly. The review includes studies of young people who have experienced abuse and are between the ages of 11 and 18. Three studies have an upper age range of older than 18. These are Buddeberg et al. (1996) – age 14–19 years; Jones et al. (1992) – age 13–19 years and Rey Gex et al. (1998) – age 15–20 years. However, the findings from these studies have been included given that there is substantial overlap in age with our population of interest. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to recognition.	of critical appraisal for observational studies are less well developed. We have therefore included this study, and not 'marked down' the overall quality rating on this basis. **Overall assessment of external validity:** + Not all studies are exactly the correct target age group. However, this only concerns 3 studies, and the overlap with our age group is substantial. **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Does the study have a UK perspective? No. Studies included in the review are from the USA (5 studies), Switzerland (2 studies), France (1 study) and New Zealand (1 study).	

5. **Evans SE, Davies C, DiLillo D (2008) Exposure to domestic violence: a meta-analysis of child and adolescent outcomes. Aggression and Violent Behavior 13: 131–40**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. Study explores the relationship between exposure to domestic violence and outcomes. **Adequate description of methodology?** Partly adequate. Good description of methods for meta-analysis. Multiple effect sizes were used for each study, rather than aggregation at the study level, which is positive given the variety of scales used. Windsorizing procedure used to recode extreme values, and test of homogeneity found no significant heterogeneity across studies. However, little information provided about included studies, including design. It is unclear why this is not reported.	**Inclusion of relevant individual studies?** Somewhat relevant. There is little data reported about the included studies. In particular, it is unclear what study design included studies used, and whether these included comparisons with non-maltreated children. However, given that odds ratios are reported, it is assumed that this must have been the case. **Study quality assessed and reported?** No. Study quality does not appear to have been assessed. However, this was a common feature across the systematic reviews of observational data which we found, and there is recognition that methods of critical appraisal for observational studies are less well developed. We have therefore included this study, and	**Does the study's research question match the review question?** Yes. Study explores the relationship between exposure to domestic violence and outcomes. **Has the study dealt appropriately with any ethical concerns?** Not reported. No discussion of ethical issues. **Were service users involved in the study?** No. Service users involved as participants in the included studies, but not included in systematic review itself. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups**	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + Good methodological detail regarding meta-analysis, and approach appears sound, however little information regarding included studies. The review does not critically appraise included studies. However, this was a common feature across the systematic reviews of observational data which we found, and there is recognition that methods of critical appraisal for observational studies

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Rigorous literature search? Yes. Nine bibliographic databases searched, reference harvesting from other SRs and citation searching.	not 'marked down' the overall quality rating on this basis. **Do conclusions match findings?** Yes. Conclusions match meta-analysis findings, and are also compared with results of previous systematic reviews.	**covered by the guideline?** Yes. Children and young people under the age of 18 who have been exposed to domestic violence. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Relates to recognition. **Does the study have a UK perspective?** Unclear. Locations where included studies were conducted is not reported.	are less well developed. We have therefore included this study, and not 'marked down' the overall quality rating on this basis.

6. **Gilbert AL, Bauer NS, Carroll AE et al. (2013) Child exposure to parental violence and psychological distress associated with delayed milestones. Pediatrics 132: e1577–83**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: 'To examine the association between parental report of intimate partner violence (IPV) and parental psychological distress (PPD) with child attainment of developmental milestones' (pe1577). This data extraction focuses on the association between intimate partner violence and language development milestones.	**Measurements and outcomes clear?** Partly. It is unclear whether, for participants who identified as Spanish speaking (21.5%) child language development was assessed in English or Spanish. **Measurements valid?** Unclear. See query regarding assessment of Spanish speakers. Also, presence of intimate partner violence	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Partly. Study approved by Indiana University Office of Research Administration. However, no mention in article about how consent was	**Overall assessment of internal validity:** - **Overall assessment of external validity:** - **Overall validity rating:** - Although associations between

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: Cross-sectional study. **Objectives of study clearly stated?** Yes. To examine the association between parental-reported intimate partner violence and parent psychological distress with children's attainment of developmental milestones. **Clearly specified and appropriate research design?** Unclear. Study is described as 'cross-sectional' but in fact has a longitudinal element, as children's milestones were monitored up until the age of 72 months. It is unclear whether, to be included in the study, children had to have 72 months' worth of data. **Subjects recruited in acceptable way?** Unclear. Study is an analysis of routinely collected health data. It is unclear whether and how participants gave consent for data to be used in this way. **Sample representative of defined population?** Unclear. The authors do not evaluate the extent to which the study sample is representative of the local population.	was determined by parental self-report and could possibly have been under-estimated. The study acknowledges that the rates of reported intimate partner violence (2.5%) is lower than has been reported in other similar settings. **Setting for data collection justified?** Yes. **Are all important outcomes and results considered?** Yes. **Tables/graphs adequately labelled and understandable?** Yes. **Appropriate choice and use of statistical methods?** Unclear. Multivariate logistic regression modelling, adjusting for parental report of child abuse concern, sociodemographic characteristics, clinic, language and insurance type. However, given that it appears that each participant may have multiple data points relating to different times (and different participants may have different numbers of data points) it is unclear whether regression was conducted using individual-level data, or per milestone (meaning that numerous data points from the same	obtained from participants, particularly as study uses routinely collected health data. **Were service users involved in the study?** No. Service users involved as participants, but not in design or interpretation of results. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children aged up to 72 months who have been exposed to intimate partner violence within their family. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to recognition. **Does the study have a UK perspective?** No. The study was conducted in the USA (Indiana). Potential impact in that methodology suggests that there are a high number of Spanish speakers in	exposure to intimate partner violence and language milestones were adjusted for language, it is a concern that it is unclear whether there was an option to assess language milestones in Spanish as well as English, given that 21.5% of participants identified as Spanish-speaking. If Spanish speakers are over-represented amongst those who have been exposed to intimate partner violence (also unclear) this could have artificially inflated the association between exposure and missed language milestones.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	individual could be used in the same regression). If the latter is the case, older children will have a greater influence on the regression calculate, as will represent a greater number of data points. **In-depth description of the analysis process?** Yes. **Are sufficient data presented to support the findings?** Partly. Regression coefficients for all included variables not reported. **Results discussed in relation to existing knowledge on the subject and study objectives?** Yes. **Results can be generalised?** Unclear. **Do conclusions match findings?** Yes.	this area – it is unclear whether this has been taken in to account when assessing the language abilities of children, i.e. whether language test is in Spanish or English.	

7. **Govindshenoy M, Spencer N (2006) Abuse of the disabled child: A systematic review of population-based studies. Child: Care, Health and Development 33: 552–8**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: To '… ascertain the strength of the association between childhood disability and abuse and neglect' (p552). **Methodology:** Systematic review.	**Inclusion of relevant individual studies?** Yes. The included studies are relevant to the topic and objectives of the review itself as well as the NCCSC review, however the Guideline Committee	**Does the study's research question match the review question?** Yes. The study aimed to '… ascertain the strength of the association	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. The review's objectives are clear and these are relevant to the NCCSC review. The authors provide an acceptable level of detail in relation to search strategies and the inclusion/exclusion criteria used. **Adequate description of methodology?** Yes. The detail provided in relation to search strategies, screening and data extraction and analysis are comprehensive. Odds ratios and confidence intervals were calculated from available data if not already reported by the individual study. **Rigorous literature search?** Yes. An appropriate number of relevant databases were searched, however the authors do not specify whether controlled vocabulary or free text terms were used. Hand searching of key journals and citation searching was carried out.	may wish to note that the 4 included represent only a small subset of the types of disabilities which the authors note in their preliminary discussion. **Study quality assessed and reported?** Yes. The quality of the included studies was assessed using an established framework. Studies were scored on a number of criteria such as sample size, attrition rate, confounding variables accounted for, definition of disability, definition of abuse. These were then summed to give a total score out of 8. **Do conclusions match findings?** Yes.	between childhood disability and abuse and neglect' (p552). **Has the study dealt appropriately with any ethical concerns?** No. The review authors do not record the consent processes of the individual studies or whether research protocols were approved by institutional review boards. **Were service users involved in the study?** Not reported. The review does not record whether service users were involved at the design stage or in the interpretation of results for any of the included studies. **Is there a clear focus on the guideline topic?** Yes. The study focuses on the association between disability and abuse and neglect in childhood. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. The samples of the reported studies include children and young people with experience of abuse and neglect.	++ **Overall validity rating:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Is the study setting the same as at least 1 of the settings covered by the guideline? Partly. No details provided. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Recognition - indicators relating to children and young people. **Does the study have a UK perspective?** Unclear. Two of the 4 included studies were conducted in the United Kingdom and the review itself was carried out by researchers based in England; however the two other included studies were conducted in the USA and Chile.	

8. **Hindley N, Ramchandani PG, Jones DPH (2006) Risk factors for recurrence of maltreatment: A systematic review. Archives of Disease in Childhood 91: 744–52**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: To '… systematically review the research base predicting those children at highest risk of recurrent maltreatment' (p744). **Methodology:** Systematic review. **Appropriate and clearly focused question?** Yes. Recognition/assessment of risk factors.	**Inclusion of relevant individual studies?** Yes. Inclusion criteria - cohort studies mostly retrospective. **Study quality assessed and reported?** Yes. Quality of studies assessed by 2 reviewers, using a scoring system adapted from Altman 2001.	**Does the study's research question match the review question?** Yes. Risk factors. **Has the study dealt appropriately with any ethical concerns?** Not reported. **Were service users involved in the study?** Not reported.	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** ++ **Overall validity rating:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Adequate description of methodology? Yes. **Rigorous literature search?** Yes. Electronic databases, reference checking, hand searching, personal communication.	**Do conclusions match findings?** Yes.	**Is there a clear focus on the guideline topic?** Yes. Relating to recognition/assessment of child abuse and neglect. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Partly. Not reported. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Recognition/assessment. **Does the study have a UK perspective?** No. Included studies from different countries. Most were from the USA and Australia.	Well conducted.

9. **Jones L, Bellis MA, Wood S et al. (2012) Prevalence and risk of violence against children with disabilities: A systematic review and meta-analysis of observational studies. Lancet 380: 899–907**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: To '… synthesise evidence for the prevalence and risk of violence against children with disabilities' (p899). The review reports on a meta-analysis of studies in which the prevalence or risk	**Inclusion of relevant individual studies?** Yes. The included studies are clearly relevant to the objectives of the review and to the NCCSC guideline, however it should be noted that studies were	**Does the study's research question match the review question?** Yes. The study aimed to '… syn-	**Overall assessment of internal validity:** ++ Although this appears to be a well-

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
of violence was reported in children with disabilities only, and those in which prevalence or risk in disabled children was compared to prevalence or risk in non-disabled children. As the NCCSC has focused on comparative studies in relation to questions on recognition only data from studies with a non-disabled comparison group has been reported by the NCCSC. **Methodology:** Systematic review. Included a meta-analysis. **Appropriate and clearly focused question?** Yes. The objective of the review is clear and has relevance to the NCCSC review work. Details on the search strategy and inclusion/exclusion criteria are comprehensive. **Adequate description of methodology?** Yes. The detail provided in relation to search strategies, screening, and data extraction and analysis are generally comprehensive. Odds ratios were extracted from original studies. If these were not reported the review authors calculated these if the necessary raw data were available. Random effects pooled odds ratios were then calculated	excluded if the response rate was below 50%, or if the response rate was not reported. **Study quality assessed and reported?** Yes. The authors assessed the quality of each study by scoring in relation to a range of criteria such as sample size, measures used to determine experience of violence and disability status, descriptions of participants (including the comparison group), whether odds ratios and confidence intervals were reported, etc. The authors also report that they excluded 2 studies involving children who were deaf and also had a primary diagnosis of substance use disorder '… because of the strong association between these disorders and violence' (p901). **Do conclusions match findings?** Yes.	thesise evidence for the prevalence and risk of violence against children with disabilities' (p899). **Has the study dealt appropriately with any ethical concerns?** Not reported. The review authors do not record the consent processes of the individual studies or whether research protocols were approved by review boards. **Were service users involved in the study?** Not reported. The review does not record whether service users were involved at the design stage or in the interpretation of results for any of the included studies. **Is there a clear focus on the guideline topic?** Yes. The study aims to determine the association between disability in children and young people and abuse. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. The review aimed to determine the extent to which children and young people with a disability are at risk of being abused.	conducted review and meta-analysis it should be noted that very little detail is provided in relation to the characteristics of participants in the individual studies, and in some cases the information that is provided suggest that the definition of abuse used by each study may be quite wide. Findings should therefore be used with caution. **Overall assessment of external validity:** ++ **Overall validity rating:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
and heterogeneity and risk of bias were estimated using the I2 statistic. Detail in relation to whether disability was clinically diagnosed and how experience of violence was reported is minimal as the review authors simply state whether the study met or did not meet the quality score criteria in relation to these measurements. **Rigorous literature search?** Yes. An appropriate number of relevant databases were searched, and both controlled vocabulary and free text were used. Hand searching of key journals and citation searching were also conducted.		**Is the study setting the same as at least 1 of the settings covered by the guideline?** Partly. Not reported. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Recognition – indicators relating to children and young people. **Does the study have a UK perspective?** Unclear. The majority of included studies were conducted in the USA, however 2 of those applicable to question 3 of the NCCSC review were conducted in the UK. Researchers based in the UK also participated in the review process.	

10. Kočovská E, Puckering C, Follan M et al. (2012) Neurodevelopmental problems in maltreated children referred with indiscriminate friendliness. Research in Developmental Disabilities 33: 1560–5

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: The study aimed to '… explore the extent of neurodevelopmental difficulties in severely maltreated adopted children' (p1560). We have extracted only data relating to language. **Methodology:** Cross-sectional study.	**Measurements and outcomes clear?** Yes. The study focused on neurodevelopmental difficulties and psychiatric disorders such as intelligence and post-traumatic stress disorder, and the scales used to measure these outcomes are reported.	**Does the study's research question match the review question?** Yes. The study aimed to '… explore the extent of neurodevelopmental difficulties in severely maltreated adopted children' (p1560). We have extracted only data relating to language.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Objectives of study clearly stated? Yes. **Clearly specified and appropriate research design?** Partly. The authors do not discuss the research design in detail but the methodology (observational study comparing a group of adopted children with experience of severe maltreatment children to a group of non-maltreated children) was appropriate to the aim of the study. **Subjects recruited in acceptable way?** Partly. Adopted children were recruited through an adoption charity, which approached eligible families living within travelling distance of the clinic. The researchers aimed to focus on children with early experience of maltreatment who were now living in a stable environment and this was ensured through clear exclusion criteria in relation to this. The comparison group were recruited through two GP surgeries, each with more than 750 children registered between the ages of 5 and 12. The authors report that this group was not intended to be '… a	**Measurements valid?** Partly. All assessments were conducted using pre-established measures, however data in relation to reliability and validity of these are not reported. It should also be noted that the Manchester Child Attachment Story Task (Green et al. 2000, data not extracted as this does not relate to language ability) was designed for use with children up to the age of 8; however the study included children up to the age of 12 (mean age of adopted group = 9.4 years, mean age of non-adopted group = 8.7 years). Pre-adoption histories (i.e. in relation to history of maltreatment, birth weight, etc.) of the adopted group were taken from social worker notes using a checklist designed for the study. **Setting for data collection justified?** Yes. It should be noted that whilst many of the tests that the adopted group completed were conducted in a clinic setting, those completed by the comparison group were conducted in general practitioner's surgeries or family homes to enable participation. **Are all important outcomes and results considered?** Yes.	**Has the study dealt appropriately with any ethical concerns?** Yes. A regional ethics committee approved the protocol and consent was sought from parents and children. **Were service users involved in the study?** No. No indication that service users were involved in the design of the study or interpretation of findings. **Is there a clear focus on the guideline topic?** Yes. The study focuses on the neurodevelopmental consequences of severe maltreatment. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. The study focused on neurodevelopmental difficulties exhibited by adopted children who had experienced severe maltreatment. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Data were collected in participant's homes, GP surgeries, or a clinic setting.	Little justification given for choice of statistical tests.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
representative sample of the general population but, rather, to achieve a group of typically developing children, matched on age and gender with the adopted group' (p1561). Due to imbalances in age and gender, a second set of recruitment letters was sent to families with boys between the ages of 6 and 10. Children were also excluded from the comparison group if they had a psychiatric disorder however it appears that some children in the comparison group were assessed as having a possible psychiatric disorder by the researchers. The authors do not describe how these two practices were selected or the demographic characteristics of the areas in which they were located or the socioeconomic status of participants, which may also be important when considering language ability specifically. **Sample representative of defined population?** Unclear. The authors do not compare recruited children to the wider population of adopted children with experience of severe maltreatment.	**Tables/graphs adequately labelled and understandable?** Yes. **Appropriate choice and use of statistical methods?** Yes. t tests used to examine between-group differences for continuous variables, chi-square and Fisher's exact test used for categorical variables. The authors do not appear to have corrected for multiple statistical comparisons. **In-depth description of the analysis process?** No. The authors simply report the statistical tests used for results in relation to each measure. No justification is given for choice of statistical tests. **Are sufficient data presented to support the findings?** Yes. **Results discussed in relation to existing knowledge on the subject and study objectives?** Yes. **Results can be generalised?** Unclear. **Do conclusions match findings?** Yes.	**Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Recognition – indicators relating to children and young people. **Does the study have a UK perspective?** Yes. The study was conducted in Scotland.	

11. Lereya ST, Samara M, Wolke D (2013) Parenting behavior and the risk of becoming a victim and a bully/victim: a meta-analysis study. Child Abuse & Neglect 37: 1091–108

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. Study examines the association between parenting behaviour and risk of becoming a victim or bully/victim. **Adequate description of methodology?** Yes. Methods for meta-analysis well described. Random effects model used for analysis and distribution examined using tests of heterogeneity. Publication bias also accounted for using Rosenthal's failsafe method (Rosenthal, 1979). **Rigorous literature search?** Partly rigorous. Four bibliographic databases were searched, but there were no hand searches.	**Inclusion of relevant individual studies?** Yes. **Study quality assessed and reported?** No. The review does not critically appraise included studies. However, this was a common feature across the systematic reviews of observational data which we found, and there is recognition that methods of critical appraisal for observational studies are less well developed. We have therefore included this study, and not 'marked down' the overall quality rating on this basis. **Do conclusions match findings?** Yes. Narrative findings are in accordance with data reported in tables.	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Not reported. No ethical considerations reported. **Were service users involved in the study?** No. Service users involved as participants in the included studies, but not included in systematic review itself. **Is there a clear focus on the guideline topic?** Yes. Relevant to question on recognition. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children and young people experiencing abuse or neglect aged between 4 and '12+'. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + Key limitations: no critical appraisal of included studies. However, the rest of the systematic review is of high quality, and statistical data is well reported. This has therefore been rated as 'moderate' quality.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Does the study have a UK perspective? No. Of the 6 studies relevant to abuse and neglect, 1 is US, 4 are categorised as 'Europe' and 1 is categorised as 'Other'.	

12. Luke N, Banerjee R (2013) Differentiated associations between childhood maltreatment experiences and social understanding: A meta-analysis and systematic review. Developmental Review 33: 1–28

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. The review question is clearly focused and is relevant to the NCCSC review work; and there is a good level of detail provided in relation to the searches used and the inclusion/exclusion criteria. **Adequate description of methodology?** Yes. There is a good level of detail provided in relation to the methodology used in the meta-analysis (random effects model). Cohen's d effect sizes (weighted by sample size) are used and significance levels are reported. For studies included in the systematic review, there are no quantitative details reported and these findings have therefore not been extracted by the NCCSC. NB – effect sizes for those studies included in the	**Inclusion of relevant individual studies?** Somewhat relevant. The included studies are on the whole appropriate to address the review question as set out by the authors and are clearly relevant to the NCCSC work; however it is not clear why a small number of studies with an adult sample were included when the review seeks to determine the impact of maltreatment on the social understanding of children. Data extracted by the NCCSC are taken from the meta-analysis in which one of the 19 included studies had an adult sample (100%). In addition, it should be noted that studies which focused only on children and young people who had experienced sexual abuse (rather than those comparing sexual abuse to other subtypes of abuse/maltreatment) were excluded.	**Does the study's research question match the review question?** Yes. The objective of the review is to '… evaluate the strength of evidence for the hypothesis that physically abused or neglected children underperform relative to their nonmaltreated peers in measures of social understanding' (p2). **Has the study dealt appropriately with any ethical concerns?** Not reported. The review authors do not record whether the research protocols of individual studies were approved by institutional review boards or how consent was dealt with. **Were service users involved in the study?** Not reported. The review does not record whether service users were involved at the	**Overall assessment of internal validity:** + The inclusion of studies in which the sample was partially or wholly comprised of adults (and the lack of discussion in relation to this decision), the lack of information on the quality of included studies, and the small number of databases searched are areas of concern. **Overall assessment of external validity:** ++ **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
meta-analysis have also been reported separately by the NCCSC. **Rigorous literature search?** Partly rigorous. The authors only searched 2 databases, although citation searching and hand searching of key journals was also carried out. The search terms used appear to be comprehensive and relevant however the full search strategy is not reported, and it is not clear whether controlled vocabulary was used. The search excluded articles containing 'review' as a keyword, which may have been a means of excluding systematic reviews, however this is not explained by the authors.	**Study quality assessed and reported?** Partly reported. The authors report that they made notes in relation to sample (e.g. representativeness), recruitment, validity of measures, choice of statistical analysis, and the reported conclusions; however these do not appear to have been translated into a formal quality rating and these notes are not reported in the review. **Do conclusions match findings?** Yes.	design stage or in the interpretation of results for any of the included studies. **Is there a clear focus on the guideline topic?** Yes. The study focuses on the impact of maltreatment on the social understanding of children. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. The review focuses on children who have experienced maltreatment. **Is the study setting the same as at least 1 of the settings covered by the guideline?** No. Not reported. The review authors do not report whether the individual studies recorded their settings or contexts. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Recognition - indicators relating to children and young people. **Does the study have a UK perspective?** No. The majority of studies were conducted in the	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		USA (17/19). The review was conducted by researchers based in the UK.	

13. Miller AB Esposito-Smythers C, Weismoore JT et al. (2013) The relation between child maltreatment and adolescent suicidal behavior: A systematic review and critical examination of the literature. Clinical Child and Family Psychology Review 16: 146–72

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. Examines the relationship between child maltreatment and adolescent suicidal ideation and attempts. **Adequate description of methodology?** Partly adequate Limited report of quality assessment of included studies, and no reporting of statistical data for any studies, making it difficult to verify the conclusions drawn. The authors do not justify why no statistical data are reported, or statistical analyses conducted. **Rigorous literature search?** Partly rigorous. Three bibliographic databases searched, with reference harvesting from included studies and other major reviews. However, no hand searching and relatively restricted range of bibliographic databases.	**Inclusion of relevant individual studies?** Yes. **Study quality assessed and reported?** Partly reported. The author did not provide explicit information on quality assessment of included studies, which are mostly observational studies. However, the authors discuss and acknowledge the methodological limitations of the included studies (retrospective longitudinal and cross-sectional) with their inherent bias of retrospective design and issues of self-report and recall bias. The authors suggest caution in the interpretation of the evidence findings. However, the overall direction of the evidence appears to show similar patterns emerging from both study designs in some areas. **Do conclusions match findings?** Yes.	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Not reported. **Were service users involved in the study?** Not reported. Service users involved as participants in the included studies, but not included in systematic review itself. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + Of moderate quality in terms of a narrative review of studies. Unclear why statistical data not reported, or statistical analyses not conducted.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Does the study relate to at least 1 of the activities covered by the guideline? Yes. **Does the study have a UK perspective?** No. Systematic review of studies from different countries: USA (13 studies), New Zealand (4 studies), Switzerland (2 studies), Canada (2 studies), Brazil (1 study), Netherlands (1 study), Italy (1 study), Australia (1 study), France (1 study), country not reported (26 studies).	

14. Mironova P, Rhodes AE, Bethell JM, et al. (2011) Childhood physical abuse and suicide-related behavior: A systematic review. Vulnerable Children and Youth Studies 6: 1–7

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. Study examines relationship between childhood physical abuse where the perpetrator is identified as a family member or parent and suicide-related behaviour. **Adequate description of methodology?** Yes. Good description of data extraction, including process for reviewing disagreements. The authors state that results were not pooled meta-analytically, as there was a small number of	**Inclusion of relevant individual studies?** Yes. **Study quality assessed and reported?** Partly reported. Formal quality assessment rules were not applied, given the lack of consensus and evaluation tools to assess observational studies (Sanderson et al. 2007). **Do conclusions match findings?** Yes.	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Not reported. **Were service users involved in the study?** Not reported. **Is there a clear focus on the guideline topic?** Yes.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** + Good reporting of methodology, and justification for lack of critical appraisal of included studies.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
studies with methodological heterogeneity across studies. **Rigorous literature search?** Yes.		**Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. School and population setting. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. **Does the study have a UK perspective?** No. One study each from South Africa, Hong Kong, USA, New Zealand, Canada.	

15. Naughton AM, Maguire SA, Mann MK et al. (2013) Emotional, behavioral, and developmental features indicative of neglect or emotional abuse in preschool children: a systematic review. JAMA Pediatrics 167: 769–75

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. Objective: 'To define the emotional, behavioural and developmental features of neglect or emotional abuse in preschoolers' (p769). **Adequate description of methodology?** Partly adequate. Methodological detail provided in a series of appendices. However, it is unclear how studies have been synthesised to arrive at the lists of	**Inclusion of relevant individual studies?** Yes. **Study quality assessed and reported?** Partly reported. Study quality was assessed using a series of critical appraisal checklists, depending on study design. Critical appraisal included consideration such as whether researchers rating behaviours etc. were 'blind' to condition, sampling procedures and so on. The quality standard	**Does the study's research question match the review question?** Yes. Study focuses on indicators of neglect and emotional abuse. **Has the study dealt appropriately with any ethical concerns?** Yes. Systematic review – no particular ethical issues.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall validity rating:** - Key limitations: critical appraisal conducted but unclear how this

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
indicators reported in Tables 1 and 2. There is poor reporting of statistical data from the included studies. **Rigorous literature search?** Yes. Eighteen bibliographic databases searched, 2 journals hand searched (Child Abuse and Neglect; Child Abuse Review), 6 websites searched.	assigned to each study is reported in an appendix, however it is not clear how this was brought to bear on the analysis. **Do conclusions match findings?** Partly. The main findings are 2 tables summarising the indicators of neglect and emotional abuse. It is unclear how these have been derived from the included studies.	**Were service users involved in the study?** No. Service users involved as participants in the included studies, but not included in systematic review itself. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Included studies concerned children aged 0–6 years experiencing neglect, emotional abuse or emotional neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to recognition. **Does the study have a UK perspective?** No. Forty of the included studies are from the USA, 2 are from Canada.	was used within the analysis, unclear how study results were combined to arrive at the lists of indicators presented, poor reporting of statistical data from the original studies.

16. **Nolin P, Ethier L (2007) Using neuropsychological profiles to classify neglected children with or without physical abuse. Child Abuse and Neglect 31: 631–43**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: Aim of the study is to 1) investigate whether cognitive functions can differentiate neglected children with or without physical abuse compared to comparison participants; 2) demonstrate detrimental impact of maltreatment on children. Study looks at a range of cognitive functions. Only data in relation to language development have been extracted here. **Methodology:** Cross-sectional study. Comparative observational study comparing cognitive functioning of children who have experienced neglect/neglect and physical abuse with non-abused children. **Objectives of study clearly stated?** Yes. **Clearly specified and appropriate research design?** Yes. **Subjects recruited in acceptable way?** Unclear. Process for recruiting children via Child Protection Services not specified.	**Measurements and outcomes clear?** Yes. **Measurements valid?** Partly. Language measure relates to receptive language only, with no test of productive language. **Setting for data collection justified?** Yes. **Are all important outcomes and results considered?** Yes. **Tables/graphs adequately labelled and understandable?** Yes. **Appropriate choice and use of statistical methods?** Partly. Unclear what statistical data are reported here. Study states that an initial multiple analysis of variance (MANOVA) was carried out, followed by post hoc univariate analyses. However, in results table on univariate analysis results are reported. There is no Bonferroni correction of significance levels to account for multiple statistical tests. **In-depth description of the analysis process?** No.	**Does the study's research question match the review question?** Yes. Study investigates cognitive and linguistic functioning in abused versus non-abused children. **Has the study dealt appropriately with any ethical concerns?** Yes. Study approved by university ethics committee. Consent obtained from parents and from children (>10 in writing, <10 verbally). **Were service users involved in the study?** No. Service users involved as participants but not in design or interpretation of results. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population is children who have experienced neglect with or without physical abuse, and a matched control group.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall validity rating:** - Study limitations include the fact that only receptive, and not productive, language abilities were assessed. The study also had relatively small sample size, particularly for the neglect without physical abuse subgroup.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Sample representative of defined population? Unclear. No comparison of sample with population of either abused children, or general child population.	**Are sufficient data presented to support the findings?** Yes. **Results discussed in relation to existing knowledge on the subject and study objectives?** Partly. Non-significant results not discussed in detail. **Results can be generalised?** Unclear. Validity of measure used for language is unclear, given that only receptive language was measured. **Do conclusions match findings?** Partly. Although there is little discussion of non-significant results.	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to recognition. **Does the study have a UK perspective?** No. Study conducted in Canada, with French-speaking children.	

17. Noll JG, Shenk CE, Yeh MT et al. (2010) Receptive language and educational attainment for sexually abused females. Pediatrics 126: e615–22

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: Aim of the study to '… test whether the experience of childhood sexual abuse is associated with long-term receptive language acquisition and educational attainment deficits for females …' (pe615).	**Are the outcomes clearly defined?** Yes. **Is the assessment of outcome blind to exposure status?** Unclear. No details provided. **If blinding was not possible, is there some recognition that knowledge of exposure status**	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval received by institutional board. Parental consent and child assent obtained for participants under age 18. After	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: Other – prospective cross-sequential design, following participants up until the age of 30. Only data for <18 have been extracted here. This study has been critically appraised using a prospective cohort tool. **Does the study address an appropriate and clearly focused research question?** Yes. **Are the two groups being studied from source populations that are comparable in all respects other than the factor under investigation?** Yes. The authors report that there were no statistically significant differences between the 2 groups in relation to minority or socioeconomic status (statistical data is not provided). They also report that the 2 groups were similar in relation to residence (zip codes), age, 'family constellation', and other nonsexual traumatic events however the statistical significance of between group differences in relation to these is not reported. **Does the study indicate how many how many of the people asked to take part did so (in**	**could have influenced assessment of outcome?** Unclear. **Is the assessment method for exposure reliable?** Yes. Substantiation by Child Protective Services. **Is evidence from other sources used to demonstrate that method of outcome assessment is valid and reliable?** Yes. **Is exposure level or prognostic factor assessed more than once?** Yes. **Are the main potential confounders identified and taken into account in design and analysis?** Yes. **Are confidence intervals provided?** Not applicable.	age 18 (this data not included here) participants gave their own consent. **Were service users involved in the study?** No. Service users involved as participants, but not in study design or interpretation of results. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Females with substantiated familial sexual abuse and a non-abused comparison group. Data were gathered up to the age of 30 years, but only data up to 18 are reported here. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to recognition. **Does the study have a UK perspective?** No. The study was	+ No information provided about whether assessors were blind to participant group. Study did not appear to repeat measures of socioeconomic status and other relevant factors, which may have contributed to language development over time.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
each of the groups being studied)? No. **Does the study assess whether eligible subjects have the outcome at the time of enrolment? Is this taken into account in analyses?** Yes. **What percentage of individuals or clusters recruited into each arm of the study dropped out before the study was completed?** 96%. **Does the study compare full participants and those lost to follow up by exposure status?** Not applicable.		conducted in the USA (Washington).	

18. Pears K, Fisher PA (2005) Developmental, cognitive, and neuropsychological functioning in preschool-aged foster children: Associations with prior maltreatment and placement history. Journal of Developmental and Behavioral Pediatrics 26: 112–22

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: To examine '… a range of domains (e.g. physical growth, neuropsychological function, general cognitive function, language and executive function) in young children in foster care compared to a community sample of same-aged children from comparable socioeconomic status (SES) backgrounds' (p113).	**Measurements and outcomes clear?** Yes. **Measurements valid?** Yes. Two validated instruments used to assess language ability: language domain of the NEPSY (Korkman et al. 1998) and the Preschool Language Scale 3rd Edition (Zimmerman et al. 1991).	**Does the study's research question match the review question?** Yes. Study investigates association between maltreatment and developmental delay, including language. **Has the study dealt appropriately with any ethical concerns?** No. No mention of ethical approval or consent.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: Cross-sectional study. Observational comparative study, comparing maltreated and non-maltreated children. **Objectives of study clearly stated?** Yes. **Clearly specified and appropriate research design?** Yes. Observational comparative design. **Subjects recruited in acceptable way?** Yes. Maltreated foster children recruited via local child welfare system. Matched control recruited via advertisement. **Sample representative of defined population?** Unclear. No analysis of representativeness of sample.	**Setting for data collection justified?** Yes. **Are all important outcomes and results considered?** Yes. **Tables/graphs adequately labelled and understandable?** Yes. **Appropriate choice and use of statistical methods?** Partly. Between-groups differences explored through a series of t tests, with Bonferroni correction for multiple tests. However, within-group differences for foster children in terms of characteristics of foster placement and cognitive/neuropsychological development were explored through a series of correlations. **In-depth description of the analysis process?** Partly. Not clear if assumption of normal distribution tested, or how dealt with any outliers. **Are sufficient data presented to support the findings?** Yes. **Results discussed in relation to existing knowledge on the subject and study objectives?** Yes.	**Were service users involved in the study?** No. Service users involved as participants, but not in design or interpretation of results. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Maltreated children aged 3 to 6 and a matched comparison group. **Is the study setting the same as at least one of the settings covered by 1 guideline?** Yes. **Does the study have a UK perspective?** No. The study was conducted in the USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Results can be generalised? Yes. **Do conclusions match findings?** Yes.		

19. Prasad MR, Kramer LA, Ewing-Cobbs L (2005) Cognitive and neuroimaging findings in physically abused preschoolers. Archives of Disease in Childhood 90: 82–5

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: To '… characterise the cognitive, motor and language skills of toddlers and preschoolers who had been physically abused and to obtain concurrent MRIs of the brain …' (p82). **Methodology:** Cross-sectional study. **Objectives of study clearly stated?** Yes. **Clearly specified and appropriate research design?** Yes. Observational comparative study, with comparison group matched by age, socioeconomic status, gender. **Subjects recruited in acceptable way?** Yes.	**Measurements and outcomes clear?** Yes. **Measurements valid?** Yes. Two validated scales used to measure language: sequenced Inventory of Communication Development (Hendrick et al. 1995); for children over 36 months Clinical Evaluation of Language Fundamentals (Preschool or Third Edition) (Semel et al. 1995). Reliability of instruments not reported. **Setting for data collection justified?** Yes. **Are all important outcomes and results considered?** Yes. **Tables/graphs adequately labelled and understandable?** Yes.	**Does the study's research question match the review question?** Yes. The researchers aimed to characterise the cognitive, motor and language skills of young children who have been physically abused. Only findings in relation to language will be extracted. **Has the study dealt appropriately with any ethical concerns?** Yes. Study '… approved by and conducted in accordance with the ethical guidelines of the Institutional Review Board of the University of Texas Health Science Center at Houston …' (p83). Consent obtained from parents or, where children are under 'conservatorship' of Children's Protective Services, consent obtained from agency following placement.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + Relatively small sample size. Consideration not given to generalisability of sample, given that some participants are children who have been hospitalised due to maltreatment, and so are suffering relatively severe abuse.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Sample representative of defined population? Unclear. No analysis of whether sample is representative of wider preschool population.	**Appropriate choice and use of statistical methods?** Partly. Between-groups difference appear to have been analysed using ANOVA. However, no descriptive of analytical process, including correction for multiple tests. **In-depth description of the analysis process?** No. See above. **Are sufficient data presented to support the findings?** Yes. **Results discussed in relation to existing knowledge on the subject and study objectives?** Partly. Limited discussion in relation to existing literature. **Results can be generalised?** Unclear. Participants are children who have been hospitalised due to abuse, so perhaps more severe end of spectrum?	**Were service users involved in the study?** No. Service users involved as participants, but not in design, execution or interpretation of study. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children aged 14–77 months who had been hospitalised for physical abuse. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to recognition. **Does the study have a UK perspective?** No. Conducted in Houston, Texas.	

20. **Rhodes AE et al. (2011) Sex differences in childhood sexual abuse and suicide related behaviors. Suicide & life-threatening behavior 41: 235–54**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. The objective of the review is clear and has relevance to the NCCSC review work. The authors provide an appropriate level of detail in relation to their search strategy and the inclusion/exclusion criteria which they used. **Adequate description of methodology?** Partly adequate. The reviewers do not provide detail in relation to the calculation of odds ratios. **Rigorous literature search?** Yes. An appropriate number of relevant databases were searched, and both controlled vocabulary and free text were used. However, hand searching of key journals and citation searching are not reported.	**Inclusion of relevant individual studies?** Yes. The included studies are on the whole appropriate to address the review question as set out by the authors and are clearly relevant to the NCCSC work, however it should be noted that Choquet et al. (1997), Howard and Wang (2005), King et al. (2004), Olshen et al. (2007), and Rosenberg et al. (2005) all defined childhood sexual abuse as involving intercourse, and that results from Bagley et al. (1995) are derived by measuring 'sexual abuse outside of school settings (lifetime)' (no further details provided); and that Garnefski and Arends, (1998) also includes data from respondents over the age of 18. It should also be noted that studies with a clinical sample or participants in the child welfare system were excluded '… given that sexual abuse in these samples is formally disclosed and therefore may represent a different type of exposure than reported in the general population … Furthermore, the effects of disclosure, such as being separated from the parent, may modify the association' (p237).	**Does the study's research question match the review question?** Yes. The review's objective is to update earlier systematic reviews exploring the links between childhood sexual abuse and suicide-related behaviours. There is a particular focus on whether the strength of association differs in boys and girls. **Has the study dealt appropriately with any ethical concerns?** No. The review authors do not record the consent processes of the individual studies or whether research protocols were approved by institutional review boards. **Were service users involved in the study?** Not reported. The review does not record whether service users were involved at the design stage or in the interpretation of results for any of the included studies. **Is there a clear focus on the guideline topic?** Yes. The study aims to determine the association between sexual abuse and suicide-related behaviours in children and young people and to explore	**Overall assessment of internal validity:** + As the study did not include formal quality assessment of the included studies it is not possible to award a higher quality rating. **Overall assessment of external validity:** ++ **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Study quality assessed and reported? No. The authors report that they did not adhere to formal quality assessment rules given that there is no consensus on how to appraise observational studies and that few tools exist for this purpose. The discussion and limitations sections of the review do however provide a clear account of some of the methodological concerns associated with the included studies. **Do conclusions match findings?** Yes.	whether this association differs by gender. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children and young people under the age of 18 who have experienced sexual abuse. However it should be noted that the review excluded studies with a clinical sample or participants in the child welfare system '… given that sexual abuse in these samples is formally disclosed and therefore may represent a different type of exposure than reported in the general population … Furthermore, the effects of disclosure, such as being separated from the parent, may modify the association' (p237). **Is the study setting the same as at least one of the settings covered by 1 guideline?** No. Settings and context are not reported. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Recognition - indicators relating to children and young people.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Does the study have a UK perspective? No. Only 2 of the included studies were conducted in the UK.	

21. Spratt EG, Friedenberg S, LaRosa A et al. (2012) The effects of early neglect on cognitive, language, and behavioral functioning in childhood. Psychology 3: 175–82

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: Purpose of the study is to '… compare cognitive, language and behavioral functioning of children with no history of neglect to children with early neglectful situations, specifically those who experience physical and emotional neglect from a caregiver or deprivation due to pre-adoptive placement in an international institution environment …' (p175). **Methodology:** Cross-sectional study. Observational comparative study. **Objectives of study clearly stated?** Yes. **Clearly specified and appropriate research design?** Yes. Observational comparative design, comparing US children with a history of physical or emotional ne-	**Measurements and outcomes clear?** Yes. **Measurements valid?** Yes. Outcomes measured using validated instruments, or Child Protective Services, medical, mental health or institutional records. **Setting for data collection justified?** Yes. **Are all important outcomes and results considered?** Partly. All outcome measures reported. Results of preliminary physical tests not reported. Unclear why these were conducted. **Tables/graphs adequately labelled and understandable?** Yes. **Appropriate choice and use of statistical methods?** Yes. Between-groups differences explored	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Yes. Study approved by Institutional Review Board. Informed consent given by caregivers. Unclear whether children also asked for their consent. Study also reports that child participants underwent a physical examination, including 'vital signs', head circumference, height, weight, serum sample, urine sample and saliva sample. It is unclear what the purpose of these measures were, and the results are not reported here. **Were service users involved in the study?** No. Service users involved as participants, but not in design or interpretation of results.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** + Relatively small sample size, although statistically significant results still obtained. Physical measurements of participants taken and unclear how these were used.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
glect, children adopted from international institutions and US children with no history of neglect. **Subjects recruited in acceptable way?** Yes. Children and caregivers referred by medical or mental health practitioners, or self-referred. **Sample representative of defined population?** Unclear. Authors do not investigate sample representativeness.	using a series of one-way ANCOVAs. Appears that post hoc tests exploring significant differences were also conducted, although method for this is not reported. Multiple linear regression model used to examine predictors of outcome on 5 outcome measures. **In-depth description of the analysis process?** Partly. No information on post hoc testing. No reporting of testing assumptions for use of ANCOVA. **Are sufficient data presented to support the findings?** Yes. **Results discussed in relation to existing knowledge on the subject and study objectives?** Yes. **Results can be generalised?** Partly. Relatively small sample size. No consideration of power to detect effects given in the text. However, authors do note that, despite the small sample size, a number of statistically significant results were found. **Do conclusions match findings?** Yes.	**Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Participants were children aged 3 to 10 with a history of familial neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to recognition. **Does the study have a UK perspective?** No. The study was conducted in the USA.	

22. Stith SM, Liu T, Davies LC (2009) Risk factors in child maltreatment: A meta-analytic review of the literature. Aggression and Violent Behavior 14: 13–29

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: The study uses a meta-analytic design to determine the strength of the relationship between a range of risk factors and abuse or neglect. **Methodology:** Systematic review. **Appropriate and clearly focused question?** Yes. **Adequate description of methodology?** Partly adequate. Some good reporting of meta-analytic techniques. However, it is unclear why effect sizes using both d and r have been calculated, and what the Pearson's correlations signify in relation to categorical variables (e.g. parent gender). A high number of analyses resulted in statistically significant values for the Q^w measure of homogeneity, suggesting a high degree of heterogeneity across studies. The authors have nonetheless chosen to combine these studies, and heterogeneity is taken in to account in the limitations section only. **Rigorous literature search?** Partly rigorous. Only 1 database searched (PsychInfo). However,	**Inclusion of relevant individual studies?** Yes. **Study quality assessed and reported?** Partly reported. The process for assessing study quality is reported, and average, median and mode scores across studies. The quality rating for each individual study is not reported. **Do conclusions match findings?** There is insufficient consideration of the impact of heterogeneity between the included studies on interpretation of the findings.	**Does the study's research question match the review question?** Yes. Study is looking at association between risk factors and maltreatment. **Has the study dealt appropriately with any ethical concerns?** Not reported. **Were service users involved in the study?** No. Service users involved as participants in included studies, but not in meta-analysis itself. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Partly. Review reports that included studies must'… examine the relationship between the identified risk factor and either child physical abuse or child neglect …' (p17). However, the age of the children involved is not specified. Also important to note that '… perpetrators of child maltreatment in	**Overall assessment of internal validity:** + **Overall assessment of external validity:** - **Overall validity rating:** - Search limited to one database, and keyword searching only rather than free text searching. Unclear why effect sizes using both d and r have been calculated, and what the Pearson's correlations signify in relation to categorical variables (e.g. parent gender). A high number of analyses resulted in statistically significant values for the Q^w measure of homogeneity, suggesting a high degree of heterogeneity across studies. The authors have nonetheless chosen to combine these studies, and heterogeneity is taken in to account in the limitations section only.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
reference list for each study identified through this database was searched for potential relevant studies. Limited set of search terms used, which focused on children rather than young people. Also searched for whole terms, e.g. 'child abuse' rather than child* plus abuse.		the study must be parents or in a parenting role …' (p17). **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to recognition. **Does the study have a UK perspective?** Unclear. Countries in which included studies were conducted is not reported.	

23. Tonmyr L, Thornton T, Draca J et al. (2010) A review of childhood maltreatment and adolescent substance use relationship. Current Psychiatry Reviews 6(3): 223–34

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. There is a clear and focused review question which is relevant to the NCCSC guideline and the systematic review methodology is clearly explained and includes an appropriate level of detail. **Adequate description of methodology?** Yes. **Rigorous literature search?** Partly rigorous. The authors	**Inclusion of relevant individual studies?** Yes. The studies included in the review are relevant to the research question set out by the authors and also to the work of the NCCSC; however some of the studies include a partially adult sample. The authors note that the fact that only 2 of the included studies were longitudinal is a limitation. Studies using a clinical sample were excluded.	**Does the study's research question match the review question?** Yes. The review aims to '…identify the presence of an association between child maltreatment (neglect, witnessing domestic violence, physical, sexual and emotional maltreatment) and nicotine, alcohol and/or drug use/abuse among adolescents …' (p224). **Has the study dealt appropriately with any ethical concerns?**	**Overall assessment of internal validity:** + No formal quality appraisal. **Overall assessment of external validity:** ++ **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
searched an appropriate number of relevant bibliographic databases although they do not report whether any hand searching of key journals or citation searching was carried out. It is not clear whether free text and controlled vocabulary was used although the authors note that the full search strategy is available on request.	**Study quality assessed and reported?** Partly reported. Formal quality rating of the included studies does not appear to have been carried out however the authors report the strengths and limitations of each study in table form. **Do conclusions match findings?** Yes.	Not reported. The review authors do not record whether the research protocols of individual studies were approved by institutional review boards or whether assent or consent was provided; with the exception of Logan et al. (2009), for which parental consent rates are reported. **Were service users involved in the study?** Not reported. The review does not record whether service users were involved at the design stage or in the interpretation of results for any of the included studies. **Is there a clear focus on the guideline topic?** Yes. The study focuses on the impact of maltreatment on adolescent substance misuse. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Maltreated adolescents between the ages of 12 and 18. **Is the study setting the same as at least 1 of the settings covered by the guideline?** No. Not reported. The review authors do not report whether the individual	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		studies recorded their settings or contexts. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Recognition – indicators relating to children and young people. **Does the study have a UK perspective?** No. None of the included studies were conducted in the UK. The majority were conducted in the USA.	

24. Wilson Steven R, et al. (2010) Comparing physically abused, neglected, and nonmaltreated children during interactions with their parents: A meta-analysis of observational studies. Communication Monographs 77: 540–75

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. Study compares abused, neglected and non-maltreated children's behaviour during interactions with their parents. **Adequate description of methodology?** Partly adequate. Good description of methods for meta-analysis, use of random-effects model and testing for homogeneity and calculation of failsafe n. **Rigorous literature search?** Yes. Seven bibliographic databases searched, citation checking and	**Inclusion of relevant individual studies?** Yes. **Study quality assessed and reported?** Unclear. Clear search strategy, inclusion and exclusion criteria, but no explicit details on how the quality of included studies was assessed. This was a common feature across the systematic reviews of observational data which we found, and there is recognition that methods of critical appraisal for observational studies are less well developed. We have therefore included this study, and	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Not reported. **Were service users involved in the study?** Not reported. **Is there a clear focus on the guideline topic?** Yes.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** + Clear search strategy, inclusion and exclusion criteria, and good description of methods for meta-analysis, but no explicit details on how the quality of included studies was assessed.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
hand searching of Child Abuse and Neglect journal.	not 'marked down' the overall quality rating on this basis. **Do conclusions match findings?** Yes.	**Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Physically abused and neglected children, and their parents and carers. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Home and lab settings. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Recognition indicators. **Does the study have a UK perspective?** No. USA (26 studies), Spain (2 studies), Canada (2 studies).	

Review questions 3 and 4 – Findings tables

1. **Allen B, Tussey C (2012) Can Projective Drawings Detect if a Child Experienced Sexual or Physical Abuse? A Systematic Review of the Controlled Research. Trauma, Violence, and Abuse: A Review Journal 13: 97–111**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: 'A comprehensive literature review of the controlled research to determine whether any graphic	**Participants:** Children and young people – children and/or adolescents experiencing either sexual abuse or physical abuse. **Sample characteristics:**	Projective drawings - human figure drawings – genitalia (sexually abused children): presence of genitalia in children's drawings investigated in 6 studies (Hibbard and Hartman 1990a, 1990b; Hibbard et al. 1987; Howe et al. 1987; Sidun and Rosenthal 1987; Yates et al. 1985).	**Overall assessment of internal validity:** - **Overall assessment of external validity:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
indicators (e.g., genitalia, omission of body parts) or predefined scoring system can reliability [sic] and validly discriminate abused from nonabused children' (p97). **Methodology:** Systematic review. **Country:** Range of countries. **Source of funding:** Other – no funding received.	• Age - Children and/or adolescents. • Sex - Not reported. • Ethnicity - Not reported • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Sexual and physical abuse. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** Systematic reviews - number of studies – total of 23 reports (13 reports relating to sexually abused children, and 10 on physically abused children). **Recognition indicators measured:** **Drawing** – graphic indicators (e.g., genitalia, omission of body parts) or predefined scoring system can reliably and validly discriminate abused from non-abused children.	Only 1 study found any significant results (Hibbard and Hartman 1990b), although several studies found non-significantly higher numbers of sexually abused children drew genitalia. The review concludes that, overall, available research suggests extremely small differences in the drawing of genitalia between target and control groups. The presence of genitalia in a drawing may be indicative of either emotional problems, sexual abuse, or both, but caution is needed to interpret these findings. Individual studies: Hibbard et al. 1987 (n=104): no significant difference between drawings of sexually abused and non-abused children for the presence of 5 body parts: eyes, vagina, penis, navel and anus. Review does not report statistical data, however the authors conclude that sexually abused children are 5.4 times more likely to draw genitalia. Hibbard and Hartmann (1990b) (n=194): no significant difference between drawings of sexually abused and non-abused children for breast, navel or rectum. Sexually abused children more likely to draw vagina/penis. Review does not report statistical data. Hibbard and Hartmann (1990a) (n=129): no significant difference between presence of genitalia in sexually abused versus non-abused children's drawings, although a higher proportion of the sexually abused group (3 out of 65) than the control group (0 out of 64) drew genitalia. Review does not report statistical data. Howe, Burgess and McCormack (1987) (n=36): no significant differences between sexually abused	**Overall validity rating:** - Inconsistent reporting of statistical data from original studies. No report of quality assessment of data extraction, discussion between reviewers etc.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		(n=12 adolescent runaways reporting sexual abuse) and non-abused children (n=24 runaways with no history of abuse) in regard to genitalia, breasts, or overt sexual features. Sidun and Rosenthal (1987) (n=60): no significant differences between sexually abused children with psychiatric problems and psychiatric controls in regard to drawing genitalia, breasts, or overt sexual features. Review does not report statistical data. Yates et al. (1985b) (n=35): no significant differences between sexually abused (n=18) and non-abused children (n=17) in the presentation of female and male sexual features. Review does not report statistical data. <u>Sexually related features (indicators such as hands covering the pelvic region, trouser fly, circles, wedges, and phallic-like objects)</u>: this was investigated in 3 studies (Hibbard and Harman 1990a; Howe et al. 1987; Sidun and Rosenthal 1987). One study (Sidun and Rosenthal 1987) found statistically significant differences between abused and non-abused children. Individual studies: Sidun and Rosenthal (1987) (n=60): the non-abused children (n=30, psychiatric controls) drew significantly more pictures with a trouser fly than the sexually abused (n=30 psychiatric survivors) children. Review does not report statistical data. Using 1 composite score combining circles (e.g., buttons on clothes, balls, suns), wedges, and phallic-like objects (e.g., canes, cigarettes), the sexually abused (n=30 psychiatric survivors) children drew signifi-	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		cantly more pictures of these objects than the control group. Review does not report statistical data. Using a second composite scale of indicators (hands omitted, sufficient body integration; scale: wedges/phallic/circle sexual features, breasts, shapes), a significant difference was detected in the sexually abused (n=30 psychiatric survivors) children when compared with non-abused children (n=30, psychiatric controls). (Caution: possible Type 1 error (false positive) due to the large no. of analyses). Howe, Burgess, and McCormack 1987 (n=36): there were no significant differences between sexually abused (n=12) and non-abused children (n=24) on indicators such as covering of the genital area, transparent clothing, dark lines on the clothing around the genital area, and legs being pressed together or crossed in their drawings. Review does not report statistical data. Hibbard and Hartmann, (1990a) (n=194): there were no significant differences between sexually abused (n=94) and non-abused children (n=100) in their drawings displaying transparencies or legs pressed together. Review does not report statistical data. <u>Body parts/organisation</u>: this was investigated in three studies (Hibbard and Hartman 1990a, Howe et al. 1987, Sidun & Rosenthal 1987). Overall, available controlled research did not demonstrate that the human figure drawings of sexually abused children are any more likely than control groups of normal or emotionally disturbed children to omit, display abnormal size, or poorly integrate body parts. Individual studies:	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Hibbard & Hartmann, (1990a) (n=194): there were no significant differences between sexually abused (n=94) and non-abused children (n=100) in their drawings of 16 indicators (including the omission of numerous body parts (e.g., eyes, arms, legs, feet, mouth, neck, hands], poor integration of parts, big hands, short arms, long arms, and asymmetrical limbs). Review does not report statistical data. Howe et al. (1987) (n=36): there were no significant differences between sexually abused (n=12) and non-abused children (n=24) on the likelihood of having incomplete figures, omitting body parts, or emphasising the face or hair in their drawings. Review does not report statistical data. Sidun and Rosenthal (1987) (n=60): there were no significant differences between sexually abused (n=30 psychiatric survivors) and non-abused children (n=30, psychiatric controls) in regard to any overemphasized body parts, asymmetrical or abnormal limb length, or the omission of fingers or eyes in their drawings. However, the sexually abused (n=30 psychiatric survivors) was significantly more likely than the non-abused children (n=30, psychiatric controls) to omit hands from their drawings. In addition, the control group displayed poorer body integration than the sexually abused group, which was counter to the expectation. Review does not report statistical data. (Caution: possible Type 1 error (false positive) due to the large no. of analyses). <u>Other indicators</u> Other indicators relating to sexual abuse were explored in 3 studies (Howe et al. 1987; Hibbard and Hartman 1990a; Sidun and Rosenthal, 1987). Overall, the three studies suggested that the following	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		graphic indicators do not differentiate sexually abused and non-abused children: shading, monsters, clouds, presence of teeth, slanting figure, small figure, big figure and the use of colour. Individual studies: Howe et al. (1987): sexually abused children (n=12) were more likely to draw figures with less ambiguous gender and to display a faint line quality when compared with non-abused children (n=24). Review does not report statistical data. Sidun and Rosenthal (1987): there were no significant differences between sexually abused (n=30 psychiatric survivors) and non-abused children (n=30, psychiatric controls) in regard to drawing of sexually undifferentiated figures. Review does not report statistical data. Hibbard and Hartman (1990a): reporting in review is unclear as to which indicators relate to this section. <u>Projective drawings – kinetic family drawings (for sexually abused children)</u> Designed to include motion or an activity within the picture such as asking children to draw their family, including themselves, engaged in an activity). Differences between sexually abused and non-abused children in relation to kinetic family drawings were examined in three studies (Cohen and Phelps 1985; Hackbarth et al. 1991; Piperno et al. 2007). The review authors conclude that, overall, the quality and results of these studies and the lack of consistent findings does not support the use of kinetic family drawings for determining a history of sexual abuse.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Individual studies: Piperno et al. (2007): significant differences in kinetic family drawings scores found between sexually abused (n=12 psychiatric survivors) and control groups (n=12 normal controls). However, the review authors' note that the validity of findings are 'limited' due to poor methodological quality of the study (assessors not blind, lack of interrater reliability). Cohen and Phelps (1985): significant and small differences in kinetic family drawings scores (plus using a composite score of indicators) were found between sexually abused (n=89 sexual abuse survivors) and control groups (n=77 psychiatric controls). However, review authors again note that validity of findings is limited due to poor methodological quality of the study (low interrater reliability). Hackbarth et al. (1991): the kinetic family drawings scores of sexually abused group (n=20 receiving counselling, note Table 1 suggests that the number in this group was 30) were significantly lower than those in the control group (n=30), indicating more family problems or less support. The review authors note that the mental health status of the groups could be confounders. <u>Projective drawings – human figure drawings – omission of body parts (physically abused children)</u> This was examined in three studies (Blain et al. 1981; Culbertson and Revel 1987; Prino and Peyrot 1994). The review authors conclude that '… the omission of feet may be attributable to the status of receiving mental health services as opposed to the experience of physical abuse' (p107). Available evi-	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		dence does not suggest that omitting a bodily feature from a drawing distinguishes physically abused children from their non-abused peers. Individual studies: Culbertson and Revel (1987): there were significant differences for the omission of arms and feet, but not for the omission of eyes, nose, mouth, legs, hands, body, and neck between emotionally disturbed children (n=20), physically abused children (n=20) and children with learning disability (n=20). Review does not report statistical data. It was unclear which groups differed significantly, although the authors note that the physically abused group scored the highest on each significant finding. (The review authors note that the mixing of abuse and emotional/learning disability variables in the target group confounded the study.) Prino and Peyrot (1994): there were no significant differences for the omission of hands, feet and noses between physically abused children (n=21) and non-abused children (n=21). Review does not report statistical data. Blain et al. (1981): physically abused children (n=32 psychiatric survivors) were significantly more likely to omit feet from their drawings when compared with normal controls (n=45); but no significant differences observed when comparing physically abused children (n=32 psychiatric survivors) with psychiatric controls (n=32). Review does not report statistical data. <u>Projective drawings – human figure drawings – body parts/organisation (physically abused children)</u>	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		This was examined in four studies (Blain et al. 1981; Culbertson and Revel 1987; Hjorth and Harway 1981; Prino and Peyrot 1994). The review authors conclude that, overall, poor body integration or asymmetry of limbs unlikely to be indicative of physical abuse as significant findings were not present when the control group displayed a psychiatric condition. Results pertaining to a disproportionate size of head have not been replicated and there are conflicting findings. Overall, the available evidence does not exist to support using either vacant eyes or teeth as indications of possible physical abuse. Individual studies: Hjorth and Harway (1981): the drawings of the physically abused children (n=30) demonstrated significantly more asymmetry and horizontal arm positioning than non-abused children (n=30). Review does not report statistical data. Blain et al. (1981): there were significant differences in the size of arms and legs, as well as differences for disproportionate size of head, in drawings of physically abused children (n=32 psychiatric survivors) when compared to non-psychiatric controls (n=45) but these differences were not found in drawings of abused children (n=32 psychiatric survivors) when compared to the drawings of children receiving mental health treatment (n=32 psychiatric controls). There were no significant differences for vacant eyes and teeth between the physically abused group and either of their psychiatric and non-psychiatric controls. Review does not report statistical data.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Culbertson and Revel (1987): there were no significant differences for limb asymmetry, poorly integrated parts, or body distortions in the drawings of physically abused children (n=20) and non-abused children (n=20) diagnosed with an emotional disturbance or learning disability (n=20). This study did find a significant difference when examining a disproportionate size of head. There were significant differences for vacant eyes among their three groups of subjects. Review does not report statistical data. It was unclear which groups differed significantly, although the authors note that the physically abused group scored the highest on each significant finding. Prino and Peyrot (1994): no significant difference in presence of teeth between abused group (n=21 physical abuse survivors) and control group (n=21 normal controls). Review does not report statistical data. <u>Other indicators</u> Other indicators were examined in 4 studies (Blain et al. 1981; Culbertson and Revel 1987; Hjorth and Harway 1981; Howe et al. 1987). No evidence was found to suggest that any of the following are present more often in the drawings of physically abused children: clouds, fruit on trees, person composed of geometric shapes, unusually large figures, environmental objects, and the use of colour. <u>Projective drawings – favourite kind of day</u> Favourite kind of day drawings scored for presence of inclement weather (e.g., rain, snow), a disproportionate size or excessive amount of the weather features for physically abused children compared to controls. This was examined in 3 studies (Manning	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		1987; Veltman and Browne 2000, 2001). The review authors conclude that, overall, available evidence does not warrant the use of favourite kind of day technique in detecting the physical abuse of children. Individual studies: Manning (1987): significantly higher raters' scores were reported for weather features (snow, rain in disproportionate size or excessive amount) in the physically abused group (n=10) when compared with normal controls (n=10) (13.7 scores vs. 4.2 scores, p<.001). Veltman and Browne 2000, 2001: there were no significant differences in the detection of weather features or raters' scores between physically abused children (n=6, n= 4 respectively) and normal controls (N=12, N= 23 respectively) in 2 studies. Review does not report statistical data. (Caution: possible Type 1 error (false positive) due to high false identification rate). <u>Projective drawings – kinetic drawings (physically abused children)</u> Kinetic family drawings were examined for physically abused children in 3 studies (Prino and Peyrot 1994; Piperno et al. 2007; Veltman and Browne 2003). The review authors conclude that, overall, available research does not support the interpretation of kinetic drawings as a tool in identifying physical abuse. Individual studies: Veltman and Browne (2003): significantly more indicators in kinetic drawings for 3 features (incomplete	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		figures, disproportionate size of family members, and disproportionate size of self in relation to family) in physically abused children (n=6) than normal controls (n=12). Review does not report statistical data. There were no significant differences between physically abused children and normal controls in their drawings for the omission of persons and distorted limbs. Prino and Peyrot (1994): physically abused children (n=21) were more likely than normal controls (n=21) to omit feet and noses from a kinetic group drawing, but no differences for the omission of mouth or presence of teeth. The difference is not statistically significant. Review does not report statistical data. Piperno et al. (2007): there were significant differences for each of the 4 indicators (graphic-representative immaturity, omission of subjects, body distortion, emotional proximity) between physically abused children (n=12) receiving therapy and normal controls (n=12). (Caution: possible Type 1 error (false positive) due to vague scoring criteria used). Review does not report statistical data (Piperno et al. 2007; Veltman and Browne 2003).	

2. De Bellis MD, Hooper SR, Spratt EG et al. (2009) Neuropsychological findings in childhood neglect and their relationships to pediatric post-traumatic stress disorder. Journal of the International Neuropsychological Society 15: 868–78

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The objective was to examine impact of neglect on IQ, reading, maths, fine-motor skills, language, visual-spatial, memory/learning and	**Participants:** • Children and young people – children who had experienced neglect (without sexual abuse) as defined by Department of Social Services records, and a non-neglected control group.	**Comparison of neglected versus non-neglected children** Composite of four language measures Multivariate analysis showed an overall statistically significant effect of abuse status on language. When not controlling for child IQ, this was of medium effect	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
attention/executive functions in 2 groups of non-sexually abused medically health neglect children, 1 with post-traumatic stress disorder and 1 without, and one nonmaltreated control group. Here we have extracted findings only in relation to impact on language. **Methodology:** Cross-sectional study.[2] **Country:** USA, North Carolina. **Source of funding:** Other – US National Institute of Mental Health grant.	• Caregivers and families – caregivers of child participants were also included in the study. **Sample characteristics:** • Age - Children - neglect with post-traumatic stress disorder group (all ages given in years) - mean age 8.30 (SD=2.17), age range 4.25 to 12.92; neglect without post-traumatic stress disorder group - mean age 7.19 (SD=2.36), age range 3.08 to 12.83; control group - mean age 7.77 (SD=1.83), age range 4.17 to 11.42. Caregiver ages not reported. • Sex - Children - neglect with post-traumatic stress disorder group 61.5% male; neglect without post-traumatic stress disorder group 45.5% male; control group 62.0% male. Caregiver sex not reported. • Ethnicity – Children - neglect with post-traumatic stress disorder group: 31.8% Caucasian neglect without post-traumatic stress disorder group: 53.85% Caucasian control group: 31.3% Caucasian caregiver ethnicity not reported • Religion/belief - Not reported. • Disability - Not reported.	size (F(8, 174)=4.00, p<0.001, partial eta squared=0.16). When controlling for child IQ this was of small to medium effect size (F(8,172)=2.87, p<0.01, partial eta squared=0.12). The direction of the relationship was investigated through a series of univariate analyses. <u>NEPSY phonological processing</u> Analysis of variance found no significant difference on this variable when not controlling for child IQ (F(2,90)=2.79, p>0.05, partial eta squared=0.06), nor when controlling for IQ (F(2,89)=0.53, p>0.05, partial eta squared=0.01). <u>NEPSY speeded naming</u> Analysis of variance found a significant difference between groups when not controlling for IQ (F(2,90)=7.73, p<0.001, still significant with Bonferroni correction) with medium effect size (partial eta squared = 0.15). Pairwise comparisons showed that both neglect groups showed statistically significantly poorer performance than the control group, but there was no significant difference between the neglect groups. A statistically significant difference remained, but with only small to medium effect size, after controlling for child IQ (F(2,89)=4.47, p<0.05, partial eta-squared = 0.09). <u>NEPSY comprehension</u> Analysis of variance found a significant difference between groups when not controlling for IQ (F(2,90)=8.45, p<0.001, still significant with Bonferroni correction) with medium effect size (partial eta	**Overall validity rating:** + No justification given for age range of participants, and not made clear whether results apply to this age range only. Whilst ages across the 3 groups were not statistically significantly different, there was variation in the mean age and age ranges across groups. This does not appear to be used as a covariate in the analysis. Unclear why only neglect and post-traumatic stress disorder group selected for within-group analysis, rather than all children who had experienced neglect.

[2] The term cross-sectional study is used here to denote an observational studies in which exposure and outcome are measured at the same time (that is, not longitudinal study). This is consistent with the terminology used in the NICE guideline development manual on study classification.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Hollingshead SES score neglect with post-traumatic stress disorder group - 37.91 (SD=15.54) neglect without post-traumatic stress disorder group: 38.62 (SD=15.72) control group: 39.38 (SD=15.51). • Type of abuse - Neglect, with post-traumatic stress disorder (n=22) neglect, without post-traumatic stress disorder (n=39) control group (n=45). • Looked after or adopted status - Proportion of children living with biological parents neglect with post-traumatic stress disorder group: 22.7% neglect without post-traumatic stress disorder group: 41.0% Controls: 95.5% • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - n=45. • Intervention numbers - neglect with post-traumatic stress disorder (n=22) neglect without post-traumatic stress disorder (n=39). • Sample size - Total sample size: n=106.	squared =0.16). Pairwise comparisons showed that both neglect groups showed statistically significantly poorer performance than the control group, but there was no significant difference between the neglect groups. A statistically significant difference remained, but with only small to medium effect size, after controlling for child IQ ($F(2,89) = 4.19$, $p<0.05$, partial eta-squared=0.09). <u>Peabody Picture Vocabulary Test</u> Analysis of variance found a significant difference between groups when not controlling for IQ ($F(2,90)=11.30$, $p<0.001$, still significant with Bonferroni correction) with medium to large effect size (partial eta squared =0.20). Pairwise comparisons showed that both neglect groups showed statistically significantly poorer performance than the control group, but there was no significant difference between the neglect groups. A statistically significant difference remained, but with only small to medium effect size, after controlling for child IQ ($F(2,89)=6.21$, $p<0.05$ still significant with Bonferroni correction, partial eta-squared =0.12). **<u>Within-group differences for neglected children with post-traumatic stress disorder</u>** Within-group analysis was conducted examining the association between post-traumatic stress disorder and maltreatment-related variables, and neuropsychological functioning for children with neglect and post-traumatic stress disorder. It is unclear why children experiencing neglect only were not included in this analysis. The findings for the language domain were as follows.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Recognition indicators measured: Language measured using: NEPSY (Korkman et al. 2001) phonological processing, speeded naming and comprehension subscales. Peabody Picture Vocabulary Test (Dunn et al. 1997).	<u>Total post-traumatic stress disorder symptoms</u> Significant negative association between total post-traumatic stress disorder symptoms and language ability, with large effect size (r=-0.50, p<0.05) post-traumatic stress disorder severity: non-significant negative association between post-traumatic stress disorder severity and language ability (r=-0.21, p>0.05). <u>Post-traumatic stress disorder cluster B (intrusive re-experiencing of the trauma)</u> Significant negative association between cluster B post-traumatic stress disorder symptoms and language ability, with large effect size (r=-0.57, p<0.01). <u>Post-traumatic stress disorder cluster C (persistent avoidance of stimuli associated with the trauma(s) or numbing of responsiveness)</u> Non-significant negative association between cluster C post-traumatic stress disorder symptoms and language ability (r=-0.20, p>0.05). <u>Post-traumatic stress disorder cluster D (persistent symptoms of increased physiological arousal)</u> Non-significant negative association between cluster D post-traumatic stress disorder symptoms and language ability (r=-0.17, p>0.05). <u>Failure to supervise index</u> Non-significant negative association between failure to supervise and language ability (r=-0.28, p>0.05). <u>Failure to provide index</u> Non-significant negative association between failure to provide and language ability (r=-0.16, p>0.05).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Witnessing family violence index Non-significant negative association between witnessing family violence and language ability (r=-0.42, p>0.05). Physical abuse index Non-significant negative association between physical abuse and language ability (r=-0.37, p>0.05). Emotional abuse index Non-significant negative association between emotional abuse and language ability (r=-0.38, p>0.05).	

3. Eigsti I, Cicchetti D (2004) The impact of child maltreatment on expressive syntax at 60 months. Developmental Science 7: 88–102

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The main aim of the study was to '… examine spontaneous language in a sample of maltreated children and well-matched comparison children, focusing specifically on the syntactic complexity of the children's utterances, to see whether the deficits observed in previous studies of maltreated toddlers extend to syntactic complexity in school-age children' (p92). A secondary	**Participants:** Children and young people. The study compared a group of maltreated children and their mothers to a demographically similar comparison group. All children were around the age of 5 when assessments took place. The maltreated group was recruited from mother-child dyads already enrolled in the Harvard Child Maltreatment Project. The authors report that none of the children '… had participated in earlier studies from the HCMP sample …' (p92). The onset of maltreatment had occurred before the age of 2 for all children in the maltreated group. Children in this group had been randomly selected from the active or current	The authors report that, for child language data, MANCOVA was conducted with group (maltreated vs. comparison) as the independent variable and outcome measures as dependent variable. However, for some measures gender was also examined as an independent variable. It is not always clear whether this was achieved via two-way MANCOVA with group and gender as independent variables, or sequential MANCOVAs. We have assumed throughout that 2-way MANCOVA was conducted, and highlighted where interaction effects were not reported. For some measures, 'effects' of socioeconomic status are also reported – here we have assumed that a separate analysis was conducted, although this is not specified. Maternal utterance data were analysed using MANCOVA with group (maltreated vs. comparison) as the independent variable, maternal utterance	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + Some lack of clarity in the paper regarding statistical analysis.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
goal of the study was determine if differences in maternal utterances were correlated with the syntactic development of children. **Methodology:** Cross-sectional study. **Country:** USA – Boston. **Source of funding:** • Government - Grants provided by the National Center on Child Abuse and Neglect, and the National Institute of Mental Health. • Voluntary/charity - A grant was provided by the Spunk Fund, Inc.	caseloads of the state social services department and their families are reported to be of a generally low socioeconomic status. To achieve a demographically similar comparison group notices were placed in welfare offices and stores in low-income neighbourhoods. The researchers ensured that this group had not experienced maltreatment by searching state databases (with permission from families). **Sample characteristics:** • Age - Maltreated group – mean age 57.6 (3.5 SD); comparison group - mean age 59.4 (1.8 SD); p=.10). • Sex - Maltreated group - male n=10, female n=9; comparison group – male n=7, female n=7; p =.88. • Ethnicity - Maltreated group - Caucasian n=17, African American n=2, Hispanic n=0, comparison group - Caucasian n=9, African American n=4, Hispanic n=1; p = .08. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - The authors state that families were generally of a low socioeconomic sta-	categories as the dependent variable, and child age as the covariate. Analysis of effects of socioeconomic status do not appear to have been conducted. <u>Syntactic complexity of spontaneous language (measured using the Index of productive syntax, lower scores correspond to less advanced structures)</u> - As scores are based on a similar number of utterances, these were controlled for overall frequency of utterances. Two-way MANCOVA conducted with maltreatment status and gender as independent variables, and using child age and maternal IQ as covariates. Maltreatment status: The maltreated group had significantly lower scores than the comparison group on measures of syntactic complexity; F(1, 27)=5.33, p=.03. No effect sizes reported. It was not possible to calculate effect sizes using the data provided. Gender: Girls had significantly lower scores than boys on measures of syntactic complexity; girls vs. boys, F(1, 27)=4.29, p=0.48. No effect sizes reported. It was not possible to calculate effect sizes using the data provided. Maltreatment status x gender: No significant interaction (no statistical data presented). Socioeconomic status: No significant interaction (no statistical data presented). <u>Auxiliary verbs in obligatory contexts (number of auxiliary verbs produced in obligatory contexts)</u> - Maltreatment status using child age and maternal IQ (also reported as maternal 'VIQ') as covariates: Maltreatment status did not have a significant effect on	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	tus with ratings of 4 or 5 on the Hollingshead Four-Factor Index, and there were no significant between group differences in relation to this (statistical data not presented). Current use of Aid to Families with Dependent Children - Maltreated group 17/19, comparison group 9/13; p=.15. Lifetime use of Aid to Families with Dependent Children - Maltreated group 19/19, comparison group 13/13; p value not provided, reported as non-significant. Current use of food stamps - Maltreated group 16/19, comparison group 13/13; p=.13. Highest grade attended (mother) - Maltreated group 11.1, comparison group 11.9; p=.14. Annual family income - Maltreated group $6000, comparison group $6360; p=.75. • Type of abuse - The authors report that children in the maltreated group were experiencing chronic maltreatment as onset had occurred before the age of 2 for all children in this group. Biological mothers are reported to have been named as perpetrator or co-perpetrator for all children. Maltreatment status was established using official social service records and social worker ratings using a maltreatment checklist (87-item interview, Giovannoni and Becerra, 1979).	the production of auxiliary verbs in obligatory contexts; F(1, 27)=2.998, p< 10. No effect sizes reported. It was not possible to calculate effect sizes using the data provided. Gender: Gender did not have an effect on the production of auxiliary verbs in obligatory contexts (significance not reported). Maltreatment status x gender: Analysis not conducted/reported. Socioeconomic status: Socioeconomic status did not have an effect on the production of auxiliary verbs in obligatory contexts (significance not reported). Receptive vocabulary level (measured using the Peabody Picture Vocabulary Test-Revised, lower scores correspond to poorer performance) - NB. One child in the maltreatment group failed to complete this assessment due to time constraints. Maltreatment status: The maltreated group had significantly lower scores than the comparison group on measures of receptive vocabulary; t(30)=2.16, p<.04. Reviewing team calculated effect size using reported data, which was medium to large (ES=-0.78). Gender: Girls had significantly lower scores than boys on measures of receptive vocabulary; F(1, 27)=4.95, p=.035. No effect sizes reported. It was not possible to calculate effect sizes using the data provided. Maltreatment status x gender: No significant interaction (no statistical data presented). Socioeconomic status: Analysis not conducted/reported. Maternal utterances - Number overall (length of time in minutes required for child to produce 100 utterances entered as a covariate) -	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	This was administered to social workers by a Phd level psychologist. Emotional abuse – 16 children in the maltreated group had experienced emotional abuse. Neglect only – 9 children in the maltreated group had experienced neglect only. Physical abuse – 10 children in the maltreated group had experienced physical abuse. Physical abuse and neglect – 9 children in the maltreated group had experienced physical abuse and neglect. Sexual abuse – The authors report that cases of sexual abuse were not included because '… as was common in the 1980s, they were rarely reported to the DSS' (p92). • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Maltreated group - n=19. • Comparison group - n=14. • Total sample - n=33. **Recognition indicators measured:** Language. <u>Assessment procedure</u> - Children and their mothers were observed during a 30 minute session in a playroom stocked with age-appropriate toys ('…	Maltreatment status: Mothers of children in the maltreated group produced significantly fewer utterances than those in the comparison group; F(1, 30)=5.58, p=.025. No effect sizes reported. It was not possible to calculate effect sizes using the data provided. <u>Maternal utterances - Wh-questions (length of time in minutes required for child to produce 100 utterances entered as a covariate, calculated as a proportion of overall utterances)</u> - Maltreatment status: Mothers of children in the maltreated group produced fewer wh-questions than those in the comparison group, however this difference was not significant; p>.05. No effect sizes reported. It was not possible to calculate effect sizes using the data provided. <u>Maternal utterances - Yes/no questions (length of time in minutes required for child to produce 100 utterances entered as a covariate, calculated as a proportion of overall utterances)</u> - Maltreatment status: Mothers of children in the maltreated group produced significantly fewer yes/no questions than those in the comparison group; F(1, 30)=4.50, p=.04. No effect sizes reported. It was not possible to calculate effect sizes using the data provided. <u>Maternal utterances - Multi-clause utterances (length of time in minutes required for child to produce 100 utterances entered as a covariate, calculated as a proportion of overall utterances)</u> - Maltreatment status: Mothers of children in the maltreated group produced significantly fewer multi-clause utterances than those in the comparison group; F(1, 30)=4.86, p=.04. No effect sizes reported.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	e.g. a house set, a punching doll, doll set, cook set …' p93). This session was observed through a one-way mirror. During the first and last 10 minutes, mothers were instructed not to initiate interactions with the child. In the intervening 10 minutes, mothers were asked to play with the child as she would normally. These sessions were videotaped and transcribed and coded by researchers blinded to maltreatment status. Transcription was conducted using standard guidelines (Brown and Hanlon 1970). 'For partially unintelligible or semantically opaque utterances, a gloss was transcribed and supplemented by phonetic representations of intelligible portions. An utterance was jointly defined by intonation contour and by the presence of a discernible pause between it and surrounding utterances' (p93); 10% of the video recordings were transcribed by 2 researchers. Inter-rater reliability for these (assessed word by word) was K=.90. When disagreements arose, the two researchers reviewed the recordings in order to achieve consensus – the agreed version was used in coding. The authors do not report inter-rater reliability testing for the coding stage.	It was not possible to calculate effect sizes using the data provided. <u>Maternal utterances - Imperatives (length of time in minutes required for child to produce 100 utterances entered as a covariate, calculated as a proportion of overall utterances)</u> - Maltreatment status: Mothers of children in the maltreated group produced fewer imperatives than those in the comparison group, however this difference was not significant; p>.05. No effect sizes reported. It was not possible to calculate effect sizes using the data provided. <u>Maternal utterances - Negative imperatives (length of time in minutes required for child to produce 100 utterances entered as a covariate, calculated as a proportion of overall utterances)</u> - Maltreatment status: Mothers of children in the maltreated group produced fewer negative imperatives than those in the comparison group, however this difference was not significant; p>.05. NB. Analysis of effects of gender, maltreatment status x gender, and socioeconomic status do not appear to have been conducted. No effect sizes reported. It was not possible to calculate effect sizes using the data provided. <u>Correlation between maternal language variables and child language scores</u> - A significant correlation was found between child production of auxiliary verbs in obligatory contexts and maternal multi-clause utterances; r(33)=.35, p=.045. A significant correlation was found between child production of auxiliary verbs in obligatory contexts and maternal wh-questions; r(33)=-.41, p=.017.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Syntactic complexity of spontaneous language (measured using the Index of productive syntax, lower scores correspond to less advanced structures) - The first 100 utterances made by the child were scored however a number of criteria had to be met (Brown 1973): '(1) only fully transcribed utterances were included; (2) compounds, proper names and ritualized; (3) reduplications were counted as single words (firetruck, quack-quack, night-night); (4) fillers like mmm were not included, nor were single-word routines (yeah, no, hi); (5) single-word requests for repetition (what ?) were not included; and (6) word-for-word repetitions (within five utterances) of self or mother were not included' (pp93–4). The authors' report that scoring 100 utterances produces '… a built-in control for between-child differences in talkativeness' (p94.) Utterances were scored for specific morphological and syntactic structures and 1 point was given if the utterance met requirements for a particular grammatical structure. 'A specific utterance might meet criteria for more than one structure: For example, an iron? would be scored for all of the following: (1) intonational question; (2) use of a noun; and (3) two-word	Child scores on the Peabody Picture Vocabulary Test-Revised were significantly negatively correlated with maternal 'production of demands' (no details provided on what this includes); r(33)=-.36, p=.04. Child scores on the Peabody Picture Vocabulary Test-Revised exhibited marginally significant correlations with maternal multi-clause utterances; r(33)=.31, p=.08. Maternal expansions and repetitions of child utterances '… were highly inversely correlated with child age in the comparison group, r=-.77, p<.001, but not in the maltreated group, r=.005, non-significant, indicating that non-maltreating mothers may be more responsive to child-specific factors' (p96). Maternal intelligence (measured using the Wechsler Adult Intelligence Scales – verbal and comprehension subscales) - NB. Two mothers in the maltreatment group did not complete this assessment due to time constraints. Maltreatment status: There were no significant differences in scores of verbal IQ between mothers of children in the maltreated group and those in the comparison group; t(27)=1.7, p=.10. Between group differences in number of maternal utterances with maternal verbal IQ scores were added to a repeated measures MANCOVA (using child age, verbal IQ and session length as covariates): There were no significant differences in number of maternal utterances; F(1, 26)=2.61, p=.11.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	combination of article plus noun. Subsequent utterances were also analyzed for each structure, with a maximum of two points per structure. A structure could be scored regardless of whether it was accurate according to adult norms. Thus, a child would have been credited for producing a past tense morpheme with the Utterance We wented to the store' (p94). <u>Auxiliary verbs in obligatory contexts (number of auxiliary verbs produced in obligatory contexts)</u> - The child's first 100 utterances were also examined to determine the number of occasions in which an auxiliary verb was added to a main verb to make the utterance grammatical. Analysis of this data used the ratio of number of required verbs produced in required contexts. <u>Receptive vocabulary level (measured using the Peabody Picture Vocabulary Test-Revised (Dunn and Dunn 1981); lower scores correspond to poorer performance)</u> - This is a non-verbal multiple choice test (does not require a verbal response or reading ability). <u>Maternal utterances</u> - The researchers also analysed maternal utterances as these were considered to	None of the variables relating to the characteristics of maternal utterances were correlated with maternal verbal IQ (statistical data not presented). Maternal verbal IQ scores significantly correlated with child scores on the Peabody Picture Vocabulary Test-Revised; r (31)=.41, p=.02. Maternal verbal IQ scores significantly correlated with child scores on the Index of productive syntax; r (31)=.38, p=.04.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	suggest possible explanations for between group differences in ability (it is noted that these may have been affected by the instruction not to initiate interactions during 2 sections of the session. Maternal utterances were coded in relation to a number of categories thought to be relevant to language development in children. These included: '(1) number of maternal utterances produced during the period in which the child produced 100 scorable utterances or 30 minutes, whichever was shorter; (2) Wh-questions (What about the blue one ?) produced in that time period; (3) Yes/No questions with inverted auxiliaries (Do you want to do the house now ?); (4) complex sentences with multiple propositions (verb plus arguments) falling within an utterance intonation contour, such as relative clauses (That looks like the bear that I got you for Christmas) or sentences with subordinate adverbial clauses; (5) imperatives (Do it like that) and (6) negative imperatives (Don't throw it at me). The total number of maternal utterances was controlled for time to best capture individual differences in the amount that mothers talked during the time that a child produced a standard number of utterances. The		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	other five maternal variables were calculated as proportions of total utterances' (p94). <u>Maternal intelligence (measured using the Wechsler Adult Intelligence Scales – verbal and comprehension subscales)</u> - The authors report that because of the significant correlation between maternal and child IQ, maternal IQ may serve as a proxy for that of the child. To control for this, the IQ of mothers was tested.		

4. **Evans SE, Davies C, DiLillo D (2008) Exposure to domestic violence: a meta-analysis of child and adolescent outcomes. Aggression and Violent Behavior 13: 131–40**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The study aimed to use meta-analysis to examine the relationship between exposure to domestic violence and children's internalising, externalising and trauma symptoms. **Methodology:** Meta-analysis of 60 studies. **Country:** Not reported. **Source of funding:** Not reported.	**Participants:** Children and young people - Included studies were all of children and young people aged under 18. **Sample characteristics:** • Age - All included studies were with children under age 18. Exact ages not reported. • Sex - For studies relating to internalising behaviour, 15 included male participants (total n=1697) and 14 included female participants (total n=1758). For studies relating to externalising behaviour,	**Statistical data** 1. Internalising behaviour 1.1 Overall sample Number of studies=58, total n=7602, mean effect size=0.48, (95% confidence interval 0.39 to 0.57), associated significance test, associated significance test differed significantly from zero (z=11.25, p<0.01). 1.2 Analysis by gender Boys: Number of studies=15, total n=1697, mean effect size =0.44 (z=6.39, p<0.01), no confidence intervals reported Girls: Number of studies=14., total n=1758, mean effect size=0.39(z=5.32, p<0.01), no confidence intervals reported No significant difference between effect sizes for girls versus boys (Qb(1)=0.34, p=0.56) 1.3 Analysis by age Preschool: Number of studies=15, total n=958, mean effect size=0.47 (z=5.43, p<0.01), no confidence intervals reported School age: Number of studies = 35, total n=4492,	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + Good methodological detail regarding meta-analysis, and approach appears sound, however little information regarding included studies. The review

	16 included male participants (total n=1787) and 13 included female participants (total n=1572). • Ethnicity - Not reported. • Religion/belief - Not reported. Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. Socioeconomic position - Not reported. • Type of abuse - Exposure to domestic violence. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** Systematic reviews: number of studies. The meta-analysis is based on 60 studies; 58 studies provided outcome data for internalising symptoms, total n= 7602 53 studies provided outcome data for externalising symptoms, total n=7,200. Six studies provided outcome data for trauma symptoms, total n=not reported. **Recognition indicators measured:** • Internalising and externalising behaviour. • Trauma.	mean effect size=0.51 (z=9.57, p<0.01), no confidence intervals reported Adolescent: Number of studies =7, total n=1509, mean effect size=0.51 (z=4.21, p<0.01), no confidence intervals reported No significant difference between effect sizes at different ages (Qb(2)=0.17, p=0.92) 1.4 Analysis by gender x age Preschool girls: Number of studies=2, total n=56, mean effect size =0.51 (z=1.50, p=0.13), no confidence intervals reported Preschool boys: Number of studies = 2, total n=162, mean effect size=0.53 (z=2.232, p<0.05), no confidence intervals reported No significant differences between preschool girls and boys (Qb(1)=0.00, p=0.95) School age girls: Number of studies =8, total n=837, mean effect size =0.41 (z=3.70, p<0.01), no confidence intervals reported School age boys: Number of studies=9, total n=839, mean effect size =0.51 (z=4.54, p<0.01), no confidence intervals reported No significant differences between school age girls and boys (Qb(1)=0.43, p=0.51) Adolescent girls: Number of studies=4, total n=784, mean effect size=0.38 (z=2.52, p<0.01), no confidence intervals reported Adolescent boys: Number of studies=4, total n=597, mean effect size = 0.43 (z=2.90, p<0.01), no confidence intervals reported No significant differences between adolescent girls and boys (Qb(1)=0.27, p=0.87) 1.5 Analysis by recruitment method Shelter: Number of studies=19, total n=2210, mean effect size=0.51 (z=7.60, p<0.01), no confidence intervals reported. Community: Number of studies=17, total n=2875, mean effect size=0.52 (z=7.79, p<0.01, *note table states 0.52, narrative states 0.51), no confidence intervals reported. Clinical: Number of studies=13, total n=1915, mean effect size=0.37 (z=4.35, p<0.01), no confidence intervals reported. No significant differences between recruitment groups (Qb(2)=2.20, p=0.33).	does not critically appraise included studies. However, this was a common feature across the systematic reviews of observational data which we found, and there is recognition that methods of critical appraisal for observational studies are less well developed. We have therefore included this study, and not 'marked down' the overall quality rating on this basis.

		2. Externalising behaviour 2.1 Overall sample Number of studies=53, total N=7200, mean effect size =0.47 (95% confidence interval 0.38 to 0.56), associated significance test differed significantly from zero (z=10.11, p<0.01) 2.2 Analysis by gender Boys: Number of studies=16, total n=1787, mean effect size=0.46 (z=5.89, p<0.01), no confidence intervals reported Girls: Number of studies=13, total n=1572, mean effect size=0.23 (z=2.71, p<0.01), no confidence intervals reported Mean effect sizes for boys and girls were significantly different from each other (Qb(1)=4.11, p<0.05). 2.3 Analysis by age Preschool: Number of studies=15, total n=1085, mean effect size=0.47 (z=6.02, p<0.01), no confidence intervals reported School age: Number of studies=32, total n=3919, mean effect size=0.50 (9.66, p<0.01), no confidence intervals reported Adolescent: Number of studies=7, total n=1509, mean effect size=0.40 (z=3.65, p<0.01), no confidence intervals reported No significant differences in mean effect sizes between age groups (Qb(2)=0.59, p=0.75). 2.4 Analysis by gender x age Preschool girls: Number of studies=2, total n=56, mean effect size =-0.22 (z=-0.63, p=0.52), no confidence intervals reported Preschool boys: Number of studies=3, total n=397, mean effect size = 0.35 (z=1.80, p<0.05), no confidence intervals reported Marginally significant differences between preschool girls and boys (Qb(1)=3.27, p=0.07) School age girls: Number of studies=7, total n=641, mean effect size=0.33 (z=2.56, p<0.05), no confidence intervals reported School age boys: Number of studies=9, total n=704, mean effect size=0.61 (z=4.92, p<0.01), no confidence intervals reported No significant differences between preschool girls and boys (Qb(1)=2.00, p=0.16) Adolescent girls: Number of studies=4, total n=784, mean effect size=0.18 (z=1.06, p=0.29), no confidence intervals reported	

		Adolescent boys: Number of studies=4, total n=597, mean effect size=0.40 (z=2.41, p<0.01), no confidence intervals reported No significant differences between preschool girls and boys (Qb(1)=1.14, p=0.29). 2.5 Analysis by recruitment method Shelter: Number of studies=15, total n=1511, mean effect size=0.45 (z=6.391, p<0.01), no confidence intervals reported. Community: Number of studies=17, total N=2950, mean effect size=0.47 (z=7.74, p<0.01), no confidence intervals reported. Clinical: Number of studies=14, total n=2150, mean effect size=0.43 (z=5.74, p<0.01 *note that table states that effect size is 0.43, narrative states that it is 0.42), no confidence intervals reported. No significant differences between groups (Qb(2)=0.38, p=0.83). 3. Trauma symptoms Due to small number of studies measuring trauma symptoms, only overall weighted mean was calculated. Number of studies=6, mean effect size=1.54 (95% confidence interval 0.38 to 2.71), associated significance test differed significantly from zero (z=2.61, p<0.01). **Narrative findings** 1. Internalising behaviour 1.1 Overall sample There is a significant association between childhood exposure to domestic violence, and internalising behaviours in children, with small to medium effect size (mean effect size =0.48; 95% confidence interval 0.39 to 0.57, z=11.25, p<0.01). 1.2 Analysis by gender There was no significant difference in internalising behaviour in girls exposed to domestic violence versus boys exposed to domestic violence (Qb(1)=0.34, p=0.56). 1.3 Analysis by age There was no significant difference in the association between domestic violence and internalising behaviour for different age groups (preschool, school age and adolescent) (Qb(2)=0.17, p=0.92) 1.4	

		Analysis by gender x age There were no significant differences in the association between domestic violence and internalising behaviour for children of different genders in the different age groups (Qb(1)=0.00, p=0.95; (Qb(1)=0.43, p=0.51; (Qb(1)=0.27, p=0.87) 1.5 Analysis by recruitment method There was no significant difference in effect sizes for children recruited via different settings: shelter, community and clinical (Qb(2)=2.20, p=0.33). 2. Externalising behaviour 2.1 Overall sample There is a significant association between exposure to domestic violence and externalising behaviours, with small to medium effect size (mean effect size=0.47, 95% confidence interval 0.38 to 0.56, z=10.11, p<0.01). 2.2 Analysis by gender The relationship between exposure to domestic violence and externalising behaviour is stronger in boys than it is in girls (Qb(1)=4.11, p<0.05), with a small to medium effect size for boys (mean effect size=0.46, no confidence intervals reported) and a small effect size for girls (mean effect size=0.23, no confidence intervals reported). 2.3 Analysis by age There was no significant difference in the association between domestic violence and externalising behaviour for different age groups (preschool, school age and adolescent) (Qb(2)=0.59, p=0.75). 2.4 Analysis by gender x age There were no significant differences in the association between domestic violence and externalising behaviour for children of different genders in the different age groups (Qb(1)=3.27, p=0.07; Qb(1)=2.00, p=0.16; Qb(1)=1.14, p=0.29). 2.5 Analysis by recruitment method There was no significant difference in effect sizes for children recruited via different settings: shelter, community and clinical (Qb(2)=0.38, p=0.83). 3. Trauma symptoms There was a significant association between exposure to domestic violence and trauma symptoms, with large effect size (mean effect	

		size=1.54, 95% confidence interval 0.38 to 2.71, z=2.61, p<0.01). However, it should be noted that this estimate is based on a relatively small number of studies (n=6).	

5. Evans E, Hawton K, Rodham K (2005) Suicidal phenomena and abuse in adolescents: a review of epidemiological studies. Child Abuse and Neglect 29: 45–58

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To review association between abuse and suicidal phenomena in adolescence, through systematic review of community- and school-based studies. **Methodology:** Systematic review of 9 studies. Eight studies are described as 'questionnaire' studies (Bensley et al. 1999; Buddeberg et al. 1996; Choquet and Menke 1989; Grossman et al. 1991; Jones 1992; Rey Gex et al. 1998; Wagan Borowsky 1999; Wright 1985). It is unclear if these are cross-sectional or longitudinal, although 7 are described as 'anonymous questionnaires' (all except Choquet and Menke	**Participants:** Children and young people - Included studies were those in which the majority of participants (90% of more) were aged 12 to 20. This means that two studies (Rey Gex et al. 1998; Wagman Borowsky et al. 1999) have some participants which are out of the age range of this review (>18). However, due to the quality of the systematic review, and the fact that the majority of participants in the studies met our criteria, a decision was taken to include this review. **Sample characteristics:** • Age - Bensley et al. (1999) -13 to 18 years Buddeberg et al. (1996) - 14 to 19 years (mean 16.0) Choquet and Menke (1989) - 13 to 16 years Fergusson et al. (1996) - 18 years Grossman et al. (1991) - 11 to 18 years Jones (1992) - 13 to 19 years Rey Gex et al. (1998) - 15 to 20 years Wagman Borowsky et al. (1999) - 12 to 18 years Wright (1985) - 17 to 18 years	**Statistical data** - The study looked separately at the association between suicidal phenomena and a) physical abuse and b) sexual abuse. **a) Physical abuse** Four studies examined the association between physical abuse and suicidal phenomena. Three of the 4 studies found a statistically significant relationship between physical abuse and suicidal phenomena (as evident in odds ratio confidence intervals). Grossman et al. (1991) - Association between physical abuse and suicidal phenomena: Odds ratio=1.9 (95% confidence interval 1.5 to 2.4). Jones (1992) - Significant differences in rate of suicidal thoughts and plans depending on frequency of being hit (chi-square=78.96, p<0.0001). Significant differences in rates of suicide attempts depending on frequency of being hit (chi-square=111.16, p<0.0001). Wagman Borowsky et al. (1999) - Association between physical abuse and suicide attempts - male: Odds ratio=3.26 (95% confidence interval 2.61 to 4.07). Association between physical abuse and suicide attempts - female: Odds ratio=3.5 (95% confidence interval 3.1 to 4.1) Wright (1985) - Association between physical abuse and 'seriously considered suicide': Odds ratio=1.67 (95% confidence interval 0.35 to 9.91, statistically non-significant).	**Overall assessment of internal validity:** + The review does not critically appraise included studies. However, this was a common feature across the systematic reviews of observational data which we found, and there is recognition that methods of critical appraisal for observational studies are less well developed. We have therefore included this study, and not 'marked down' the overall quality rating on this basis. **Overall assessment of external validity:** + Not all studies are exactly the correct target age group. However,

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
1989), which would suggest cross-sectional data. One study (Fergusson et al. 1996) is a longitudinal study in which data was gathered via structured interview. **Country:** Range of countries. Studies included in the review are from the USA (5 studies), Switzerland (2 studies), France (1 study) and New Zealand (1 study). **Source of funding:** Other - Community Fund and Oxfordshire Mental Healthcare Trust.	• Sex - Bensley et al. (1999) - 52.1% female; 47.9% male Buddeberg et al. (1996) - 57.4% female; 42.6% male Choquet and Menke (1989) - 44% female; 56% male Fergusson et al. (1996) - 50.5% female; 49.5% male Grossman et al. (1991) - 51% female; 49% male Jones (1992) - 49.1% female, 50.9% male Rey Gex et al. (1998) - 43.1% female; 56.9 male Wagman Borowsky et al. (1999) - 52.1% female; 47.9% male Wright (1985) - 47.3% female; 52.7% male. • Ethnicity - Bensley et al. (1999) - 2.4% Black (not Hispanic), 75.0% White (not Hispanic), 9.1% Hispanic, 5.9% Asian (or Pacific Islander), 3.2% Native American, 5.4% Other (or unknown) Buddeberg et al. (1996) - Not reported Choquet and Menke (1989) - Not reported Fergusson et al. (1996) - 86.2% European/Pakeha, 13.8% Maori/Pacific Islander Grossman et al. (1991) - 100% Native American Jones (1992) - Not reported Rey Gex et al. (1998) - Not reported Wagman Borowsky et al. (1999) - 100% Native American Wright (1985) - Not reported. • Religion/belief - Not reported. • Disability - Not reported.	**b) Sexual abuse** Six studies examined the association between sexual abuse and suicidal phenomena. All 5 studies found a statistically significant relationship between sexual abuse and suicidal phenomena (as evident in odds ratio confidence intervals). Bensley et al. (1999). Association of abuse and molestation with suicidal phenomena: Suicidal thoughts: Odds ratio=4.4 (95% confidence interval 3.1 to 6.2); Suicide plans: Odds ratio=6.8 (95% confidence interval 4.4 to 10.4). Non-injurious attempt: Odds ratio=12.0 (95% confidence interval17.9 to 18.4). Injurious attempt: Odds ratio =47.1 (95% confidence interval 23.2 to 95.3). 2) Association of molestation with suicidal phenomena: Suicidal thoughts: Odds ratio=1.9 (95% confidence interval 1.2 to 2.8). Suicide plans: Odds ratio =3.9 (95% confidence interval 2.2 to 6.7). Non-injurious attempt: Odds ratio=2.7 (95% confidence interval 1.5 to 4.8). Injurious attempt: Odds ratio=11.6 (95% confidence interval 3.2 to 42.3) Buddeberg et al. (1996) - Positive correlation found between 'suicidality' and sexual abuse (phi=0.16). Fergusson et al. (1996) - Association of sexual abuse with suicide attempts: Adjusted OR =4.8 (95% confidence interval 2.5 to 9.2). Grossman et al. (1991) - Association between sexual abuse and suicidal phenomena: Odds ratio=1.5 (95% confidence interval 1.2 to 1.9) Rey Gex et al. (1998) - Association of sexual abuse with suicide attempts (compared with those with no suicidal thoughts or behaviours): Odds ratio=1.5 (95% confidence interval 1.2 to 1.9) Wagman Borowsky et al. (1999) - Association between sexual abuse and suicide attempts - male: Odds ratio=4.7 (95% confidence interval 3.6 to 6.3). Association between sexual abuse and suicide attempts - female: Odds ratio=2.9 (95% confidence interval 2.5 to 3.3).	this only concerns 3 studies, and the overlap with our age group is substantial. **Overall validity rating:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** Systematic reviews: number of studies - 9 studies, total number of participants=38,935. **Recognition indicators measured:** Suicidal thoughts/behaviour. Studies examined indicators including suicidal thoughts, suicidal ideation, suicidal plans, suicide attempts (non-injurious and injurious). All measured by self-report.	**Narrative findings** Four studies examined the association between physical abuse and suicidal phenomena. Three of the 4 studies (Grossman et al. 1991; Jones 1992; Wagman Borowsky et al. 1999) found a statistically significant relationship between physical abuse and suicidal phenomena. Two of these studies reported odds ratios, the lowest being 1.9 (95% CI 1.5 to 2.4), and the highest 3.5 (95% CI 3.1 to 4.1). One study found higher odds ratios for females (3.5, 95% confidence interval 3.1 to 4.1) compared to males (3.26 95% confidence interval 2.61 to 4.07) (Wagman Borowsky et al. 1999). A second study (Jones et al. 1992) found a significant association between frequency of being hit and rates of suicidal thoughts and plans (chi-square=78.96, p<0.0001), and rates of suicide attempt (chi-square=111.16, p<0.0001). One study (Wright 1985) found a non-significant association between physical abuse and suicidal phenomena. Five studies examined the association between sexual abuse and suicidal phenomena (Bensley et al. 1999; Buddeberg et al. 1996; Fergusson et al. 1996; Grossman et al. 1991; Rey Gex et al. 1998; Wagman Borowsky et al. 1999). All 5 studies found that adolescents reporting a history of sexual abuse were more likely to report a history of suicidal phenomena. Three studies reported odds ratios, which ranged from 1.5 (95% confidence interval 1.2 to 1.9) to 47.1 (95% confidence interval 23.2 to 95.3). One study (Bensley et al. 1999) found that the size of the effect was greater depending on the seriousness of the abuse, that is whether the abuse was defined as ‘molestation’ or ‘sexual abuse’.	

6. **Gilbert AL, Bauer NS, Carroll AE et al. (2013) Child exposure to parental violence and psychological distress associated with delayed milestones. Pediatrics 132: e1577–83**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: 'To examine the association between parental report of intimate partner violence (IPV) and parental psychological distress (PPD) with child attainment of developmental milestones' (pe1577). This data extraction focuses on the association between intimate partner violence and language development milestones. **Methodology:** Cross-sectional study. **Country:** USA (Indiana). **Source of funding:** Other– Funded by US National Institutes of Health.	**Participants:** • Children and young people - Children aged under 72 months who have been exposed to intimate partner violence, as reported by their caregiver, compared to a non-exposed control. • Caregivers and families - Caregivers of children aged under 72 months who have been exposed to intimate partner violence, as reported by their caregiver, compared to a non-exposed control. **Sample characteristics:** • Age - Study reports participants as being 'younger than 72 months' (p578). However, unclear whether all participants had 72 months' worth of data, or whether some were younger than 72 months. • Sex - Children Male - 50.8%, female 49.0%, missing/unknown 0.2%. Caregiver sex not reported. • Ethnicity - Children Black 46.6%, Hispanic/Latino 36/8%, White 12.1%, Other 4.5%. Caregiver ethnicity not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported.	Multivariate logistic regression adjusting for parent-reported child abuse concerns, sociodemographic characteristics, clinic, language and insurance type found that, for parents who self-reported intimate partner violence and parental psychological distress, there was an increased risk of their child missing developmental milestones in language development (adjusted OR=2.1, 95% CI 1.3 to 3.3). For parents reporting intimate partner violence only, there was also an increased risk of their child missing developmental milestones in language development (adjusted OR=1.4, 95% CI 1.1 to 1.9).	**Overall assessment of internal validity:** - **Overall assessment of external validity:** - **Overall validity rating:** - Although associations between exposure to intimate partner violence and language milestones were adjusted for language, it is a concern that it is unclear whether there was an option to assess language milestones in Spanish as well as English, given that 21.5% of participants identified as Spanish-speaking. If Spanish speakers are over-represented amongst those who have been exposed to intimate partner violence (also unclear) this could have artificially inflated the asso-

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Intimate partner violence reported by 2.5% of sample (419 individuals). Intimate partner violence and parental psychological distress reported by 0.5% of sample (88 individuals). • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** n=16595. **Recognition indicators measured:** Language milestones assessed using the Denver Developmental Screening Test.		ciation between exposure and missed language milestones.

7. **Govindshenoy M, Spencer N (2006) Abuse of the disabled child: A systematic review of population-based studies. Child: Care, Health and Development 33: 552–58**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To '… ascertain the strength of the association between childhood disability and abuse and neglect' (p552).	**Participants:** Children and young people. Spencer et al. (2005) - Children born between 1983 and 2001 in one region of West Sussex. The disabilities for which the review reports odds ratios are autism, cerebral palsy, sensory disorders; and moderate or severe conduct disorder, non-conduct psychologi-	Significance has been inferred from confidence intervals or p values where provided/reported. <u>Spencer et al. (2005) (quality score 8/8)</u> – Cerebral palsy and all forms of abuse combined: A significant association was found between cerebral palsy and all forms of abuse combined (statistical data not presented, reported narratively by review authors); however this association was found to be	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** ++ **Overall validity rating:** ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Methodology: Systematic review of population based studies (2 longitudinal studies, one 1 birth cohort study and 1 cross-sectional survey). **Country:** Range of countries. The included studies were conducted in: Spencer et al. (2005) – UK – West Sussex. Sidebotham and Heron (2003) – UK – Southwest England. Vizcarra et al. (2001) – Chile – Temuco. Brown et al., (1998) - USA - 2 counties in upstate New York. The review was carried out by researchers based in England. **Source of funding:** Not reported.	cal disorder, speech/language disorders, and learning difficulties. Disability status appears to be on the basis of medical diagnosis and experience of abuse or neglect was determined by child protection registration with social services. Sidebotham and Heron (2003) - All children born in Avon between 01/04/91 and 31/12/92. Disability status was determined by parental report of developmental concerns. Experience of abuse or neglect was determined by child protection registration with social services. Vizcarra et al. (2001) - Quasi-randomly selected sample of mothers aged between 15 and 49 with a child under the age of 18 in Temuco, Chile. Disability status ('emotional problems' – no further details provided) and experience of abuse was determined by maternal report. Brown et al. (1998) - Random sample of families with a child aged between 1 and 10 in 1975 residing in 1 of 2 upstate counties in New York. The disabilities for which the review reports odds ratios are low verbal IQ, being anxious or withdrawn, and being 'handicapped' (need for special education). Disability status was determined by parental report and experience of abuse was determined by combining retrospective self-report by the child at the age of 18 and state records.	non-significant after adjusting for birthweight, gestational age, maternal age and socioeconomic status; odds ratio=1.79 (95% CI 0.96-3.35). Cerebral palsy and physical abuse: A significant association was found between cerebral palsy and physical abuse (statistical data not presented, reported narratively by review authors); and this remained significant after adjusting for birthweight, gestational age, maternal age and socioeconomic status; odds ratio=3.00 (95% CI 1.29-6.78). Cerebral palsy and neglect: A significant association was found between cerebral palsy and neglect (statistical data not presented, reported narratively by review authors); and this remained significant after adjusting for birthweight, gestational age, maternal age and socioeconomic status; odds ratio=2.71 (95% CI 1.08-6.80). Cerebral palsy and emotional or sexual abuse: Analysis of these associations was not undertaken due to small numbers. Moderate/severe conduct disorder and all forms of abuse combined: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; an association was found between conduct disorder and all forms of abuse combined; odds ratio=7.59 (95% CI 5.59-10.31; results of unadjusted analyses are not reported). Moderate/severe conduct disorder and physical abuse: After adjusting for birthweight, gestational	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Sample characteristics: • Age - Spencer et al. (2005) – Not reported by review. The study was a retrospective birth cohort study of children born between 1983 and 2001. Sidebotham and Heron (2003) - Not reported clearly by review. The study was a prospective birth cohort study of children born between 01/4/91 and 31/12/92. It appears that all data was collected before the child reached the age of 30 months. Vizcarra et al. (2001) - Not reported by review. The study was a cross-sectional survey of mothers aged 15-49 with a child under the age of 18. Brown et al. (1998) - Not reported by review. The study was a longitudinal cohort study of families with children between the ages of 1 and 10 in 1975. Data was collected in 1983, 1986, and 1991–3. • Sex - Not reported by review. • Ethnicity - Not reported by review. • Religion/belief - Not reported by review. • Disability - Spencer et al. (2005) - Study population included children with autism, cerebral palsy, sensory disorders (hearing and visual impairment); and moderate or severe conduct disorder, non-conduct psychological disorder, speech/language disorders, and learning difficulties. Disability status appears to be on the	age, maternal age and socioeconomic status; an association was found between conduct disorder and physical abuse; odds ratio=4.09 (95% CI 2.22-7.54; results of unadjusted analyses are not reported). Moderate/severe conduct disorder and neglect: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; an association was found between conduct disorder and neglect; odds ratio=8.22 (95% CI 4.76-14.18; results of unadjusted analyses are not reported). Moderate/severe conduct disorder and emotional abuse: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; an association was found between conduct disorder and emotional abuse; odds ratio=11.58 (95% CI 7.72-17.37; results of unadjusted analyses are not reported). Moderate/severe conduct disorder and sexual abuse: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; an association was found between conduct disorder and sexual abuse; odds ratio=7.65 (95% CI 3.56-16.41; results of unadjusted analyses are not reported). Moderate/severe non-conduct psychological disorder and all forms of abuse combined: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; a significant association was found between non-conduct psychological disorder and all forms of abuse combined; odds ratio=4.38 (95% CI 2.61-7.36; results of unadjusted analyses are not reported).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	basis of medical diagnosis. Sidebotham and Heron (2003) - Study population included children with parental reported developmental issues. Vizcarra et al. (2001) - Study population included children with maternal reported 'emotional problems' (no further details provided). Brown et al. (1998) - Study population included children with low verbal IQ, who were anxious or withdrawn, and those who were 'handicapped' (need for special education). Disability status was based on parental report. • Long term health condition - Not reported by review. • Sexual orientation - Not reported by review. • Socioeconomic position - Not reported by review. • Type of abuse - Spencer et al. (2005) - Experience of abuse was determined by child protection registration with social services. The study population included children who had experienced emotional abuse, neglect, physical abuse or sexual abuse. Sidebotham and Heron (2003) - Not reported by individual study. Experience of abuse was determined by child protection registration with social services. Vizcarra et al. (2001) - Experience of abuse was determined by maternal report. The study population included	Moderate/severe non-conduct psychological disorder and physical abuse: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; a significant association was found between non-conduct psychological disorder and physical abuse; odds ratio=3.06 (95% CI 1.13-8.28; results of unadjusted analyses are not reported). Moderate/severe non-conduct psychological disorder and neglect: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; a non-significant association was found between non-conduct psychological disorder and neglect; odds ratio=2.73 (95% CI 0.87-8.62; results of unadjusted analyses are not reported). Moderate/severe non-conduct psychological disorder and emotional abuse: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; a significant association was found between non-conduct psychological disorder and emotional abuse; odds ratio=8.04 (95% CI 4.22-15.30; results of unadjusted analyses are not reported). Moderate/severe non-conduct psychological disorder and sexual abuse: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; an association was found between non-conduct psychological disorder and sexual abuse, however this was non-significant; odds ratio=1.9 (95% CI 0.28-14.28; results of unadjusted analyses are not reported).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	children who had experienced mild or severe physical violence, or psychological violence (the review authors suggest that this is similar to emotional abuse). Brown et al. (1998) - Experience of abuse was determined by combining retrospective self-report by the child at the age of 18 and state records. The study population included children who had experienced neglect, physical abuse or sexual abuse. • Looked after or adopted status - Not reported by review. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported by review. **Sample size:** • The authors do not report a combined sample size and no meta-analysis was conducted due to the heterogeneity of the included studies. The total sample sizes of the included studies were - Spencer et al. (2005): n=119, 729. Sidebotham and Heron (2003): n=14,893. Vizcarra et al. (2001): n= 22 (it is not clear whether this figure relates to the total number of children or the total number of mothers). Brown et al. (1998): n=644 (it is not clear whether this figure relates to the total number of families or the total number of mothers).	Moderate/severe speech/language disorders and all forms of abuse combined: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; a significant association was found between speech or language disorders and all forms of abuse combined; odds ratio=2.96 (95% CI 2.22-3.96; results of unadjusted analyses are not reported). Moderate/severe speech/language disorders and physical abuse: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; a significant association was found between speech or language disorders and physical abuse; odds ratio=3.43 (95% CI 2.18-5.40; results of unadjusted analyses are not reported). Moderate/severe speech/language disorders and neglect: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; a significant association was found between speech or language disorders and neglect; odds ratio=3.79 (95% CI 2.35-6.10; results of unadjusted analyses are not reported). Moderate/severe speech/language disorders and emotional abuse: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; a significant association was found between speech or language disorders and emotional abuse; odds ratio=4.21 (95% CI 2.78-6.34; results of unadjusted analyses are not reported).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Systematic reviews - number of studies - n=4. **Recognition indicators measured:** Risk factors - • Spencer et al. (2005) - Disability status appears to be on the basis of medical diagnosis. Experience of abuse determined by child protection registration with social services. • Sidebotham and Heron (2003) - Disability status was determined by parental report. Experience of abuse determined by child protection registration with social services. • Vizcarra et al. (2001) – Disability status and experience of abuse were determined by parental report. • Brown et al. (1998) - Disability status was determined by parental report and experience of abuse was determined by combining retrospective self-report by the child at the age of 18 and state records.	Moderate/severe speech/language disorders and sexual abuse: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; an association was found between speech or language disorders and sexual abuse, however this was non-significant; odds ratio = 1.27 (95% CI 0.41-3.99); results of unadjusted analyses are not reported). Moderate/severe learning difficulty and all forms of abuse combined: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; a significant association was found between learning difficulty and all forms of abuse combined; odds ratio=4.69 (95% CI 3.75-5.86); results of unadjusted analyses are not reported). Moderate/severe learning difficulty and physical abuse: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; a significant association was found between learning difficulty and physical abuse; odds ratio=3.40 (95% CI 2.25-5.12; results of unadjusted analyses are not reported). Moderate/severe learning difficulty and neglect: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; a significant association was found between learning difficulty and neglect; odds ratio=5.34 (95% CI 3.68-7.23); results of unadjusted analyses are not reported. Moderate/severe learning difficulty and emotional abuse: After adjusting for birthweight, gestational	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		age, maternal age and socioeconomic status; a significant association was found between learning difficulty and emotional abuse; odds ratio=2.93 (95% CI 1.88-4.57); results of unadjusted analyses are not reported). Moderate/severe learning difficulty and sexual abuse: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; a significant association was found between learning difficulty and sexual abuse; odds ratio=6.38 (95% CI 3.81-10.68); results of unadjusted analyses are not reported). Sensory disorders and all forms of abuse combined: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; a non-significant association was found between sensory disorders and all forms of abuse combined; odds ratio=0.76 (95% CI 0.31-1.83); results of unadjusted analyses are not reported). Sensory disorders and physical abuse: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; a non-significant association was found between sensory disorders and physical abuse; odds ratio=0.44 (95% CI 0.06-3.13) results of unadjusted analyses are not reported). Sensory disorders and neglect, emotional abuse, or sexual abuse: Analysis of these associations were not undertaken due to small numbers. Autism and all forms of abuse combined: After adjusting for birthweight, gestational age, maternal age	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		and socioeconomic status; a non-significant association was found between autism and all forms of abuse combined; odds ratio=0.79 (95% CI 0.29-2.13; results of unadjusted analyses are not reported). Autism and physical abuse: After adjusting for birthweight, gestational age, maternal age and socioeconomic status; a non-significant association was found between autism and physical abuse; odds ratio=1.23 (95% CI 0.31-5.05); results of unadjusted analyses are not reported). Autism and neglect, emotional abuse, or sexual abuse: Analysis of these associations was not undertaken due to small numbers. <u>Sidebotham and Heron (2003) (quality score 8/8)</u> – Parental reported development concerns and abuse (not divided by subtype): After adjusting for hospital admissions, feeding difficulties, low birthweight, low reported positive attributes, temper tantrums, and unintended pregnancy; a significant association was found between parental reported development concerns and abuse; odds ratio=1.99 (95% CI 1.12-3.56; results of unadjusted analyses are not reported). <u>Vizcarra et al. (2001) (quality score 6/8)</u> – Parental reported emotional problems and psychological violence: A significant association was found between parental reports of emotional problems in the child and psychological violence ($p<0.002$; odds ratios not provided by authors of individual study).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Parental reported emotional problems and mild physical violence: A significant association was found between parental reports of emotional problems in the child and mild physical violence ($p<0.001$; odds ratios not provided by authors of individual study). Parental reported emotional problems and severe physical violence: The review authors report that no association was found between parental reports of emotional problems in the child and severe physical violence ($p=0.27$; odds ratios not provided by authors of individual study). <u>Brown et al. (1998) (quality score 5/8)</u> – Parental reported low verbal IQ and neglect: A significant association was found between parental reports of low verbal IQ in the child and neglect; odds ratio=2.70 (95% CI 1.26-5.74). Parental reported low verbal IQ and physical abuse: No statistically significant association was found between parental reports of low verbal IQ in the child and physical abuse (statistical data not presented, reported as non-significant by review authors). Parental reported low verbal IQ and sexual abuse: No statistically significant association was found between parental reports of low verbal IQ in the child and sexual abuse (statistical data not presented, reported as non-significant by review authors). Parental reports of the child being anxious or withdrawn and neglect: A significant association was found between parental reports of the child being	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		anxious or withdrawn and neglect; odds ratio=2.02 (95% CI 1.03-2.96). Parental reports of the child being anxious or withdrawn and physical abuse: No statistically significant association was found between parental reports of the child being anxious or withdrawn and physical abuse (statistical data not presented, reported as non-significant by review authors). Parental reports of the child being anxious or withdrawn and sexual abuse: No statistically significant association was found between parental reports of the child being anxious or withdrawn and sexual abuse (statistical data not presented, reported as non-significant by review authors). Parental reported presence of a 'handicap' and neglect: No statistically significant association was found between the presence of a 'handicap' and neglect (statistical data not presented, reported as non-significant by review authors). Parental reported presence of a 'handicap' and physical abuse: No statistically significant association was found between the presence of a 'handicap' and physical abuse (statistical data not presented, reported as non-significant by review authors). Parental reported presence of a 'handicap' and sexual abuse: A statistically significant association was found between the presence of a 'handicap' and sexual abuse; odds ratio=11.79 (95% CI 1.01-126.17).	

8. **Hindley N, Ramchandani PG, Jones DPH (2006) Risk factors for recurrence of maltreatment: A systematic review. Archives of Disease in Childhood 91: 744–52**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: A systematic review of cohort studies investigating factors associated with substantiated maltreatment recurrence in children **Methodology:** Systematic review. **Country:** Range of countries – 15 studies conducted in the USA, and 1 Australian study. **Source of funding:** Charity – Medical Research Council.	**Participants:** Children and young people. **Sample characteristics:** • Age - Under 18 years. • Sex - Not reported. • Ethnicity - Only 1 study examined ethnicity (which ethnic groups not reported). • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - 8 studies - any form of maltreatment (neglect, emotional abuse, physical abuse, sexual abuse); 7 studies - child sexual abuse; physical abuse - 1 study. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Total sample - The review involved 592,520 families and children.	The impact of type and severity of abuse on recurrence was investigated in 7 studies (DePanfilis and Zuravin 1999a, 1999b, 2002; Fluke et al. 1999; Fryer and Miyoshi 1994; Herrenkohl 1979; Murphy et al. 1992). (References of studies based on Table 1 and Table 2. NB. Please note discrepancies as the 7 studies quoted in the text did not match the 7 studies presented in Tables 1 and 2). <u>Risk factor - type of abuse</u> The review authors conclude that, overall, neglect is associated with the highest risk of future maltreatment. 1. Index abuse type not significantly associated with recurrence. Review does not report statistical data (DePanfilis and Zuravin 1999a). 2. Neglect cases had consistently higher recurrence rates across all three service statuses (closed, open, or continued), follow-up at 5 years. No statistical data reported (DePanfilis and Zuravin 1999b, study quality score 10). 3. Unclear what results were in relation to abuse type (Depanfilis and Zuravin 2002). 4. Neglect was most likely maltreatment type to recur (log rank $p<0.001$) in 9 out of 10 states in the USA, follow-up at 2 years (Fluke 1999, study quality score 9). 5. Re-victimisation rates: physical neglect 13.07%, emotional neglect 12.02%, lack of supervision 10.99%; cuts/welts/bruises 8.8%, sexual abuse 8.26%, follow-up at 4 years (Fryer 1994, study quality score 8). Unclear whether differences were statistically significant – no statistical data reported.	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** ++ **Overall validity rating:** ++ Well conducted.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Systematic reviews - number of studies - 16 cohort studies, mostly retrospective, published between 1979 and 2002. **Recognition indicators measured:** Risk factors for recurrence of maltreatment. Interim between index episode and recurrence ranged from 1 month to 6 years in 11 studies. Time until recurrence unknown in 6 studies.	6. Recurrence rates were reported to be lower for families in which gross neglect (44.4%) compared with physical abuse (54.1%), emotional abuse 21.4%; follow-up at 10 years (Herrenkohl 1979, poor study, quality score 4). 7. Unclear what results were in relation to abuse type. (Murphy et al. 2002). <u>Risk factor - severity of abuse</u> This was explored in 2 studies (Murphy 1992 and Swanston 2002). It is likely that severity of abuse could have an impact on the likelihood of recurrence. 1. More severe form of sexual abuse was related to subsequent notification for abuse/neglect (Chi-sq = 29.54, df=3, p=0.02), follow-up at 6 years (Swanston 2002, study quality score 9). 2. No association was found between severity or type of index maltreatment and a later return to court, follow-up at 3.5 years (Murphy 1992, poor study quality score 4, review does not report statistical data). <u>Risk factor - number of previous episodes of maltreatment</u> The number of previous maltreatment episodes as risk factors was examined in 9 studies, based on data presented in Tables 1 and 2 (Depanfilis and Zuravin 1999a, 2002; English 1999; Fluke 1999; Littel et al. 2002; Murphy 1992; Rittner 2002, Swanston 2002; Wood 1997). NB. Please note discrepancies that the 8 studies quoted in the text did not match the 9 studies presented in Tables 1 and 2). The authors suggest that number of previous	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		maltreatment episodes is likely to be a predictor of future maltreatment. 1. Prior history of maltreatment (such as number of prior Child Protective Services referrals) was found to be strongly associated with recurrent maltreatment, follow-up at 18 months (English 1999, study quality score 9, review does not report statistical data). 2. Subsequent maltreatment was found to be associated with families with previous chronic neglect, follow-up at 1 year ($p<0.001$) (Little et al. 2002, study quality score 9). 3. Return to court was reported to be more likely to be associated with >6 previous reports (Chi-sq = 4.9, df=1, $p<0.05$), follow-up at 3.5 years (Murphy 1992, poor study quality score 4). 4. A strong predictor of recurrent maltreatment was Child Protective Services investigation in last 5 years (Chi-sq=25.912, df=5, $p < 0.0001$), follow-up at 18 months (Rittner 2002, study quality score 10). 5. The likelihood of recurrence was found to increase after each subsequent maltreatment event, follow-up at two years (Fluke 1999, study quality score 9, review does not report statistical data). 6. The time between episodes of maltreatment was reported to shorten as number of maltreatment episodes increased, follow-up at 5 years. Review does not report statistical data. (DePanfilis and Zuravin 2001, study quality score 11). NB. This is taken from narrative summary text and is not reported in Table 2. 7. Data relating specifically to previous numbers of episodes is not reported (Swanston 2002, study quality score 9).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		8. Previously maltreated children were approximately 6 times more likely to experience recurrent maltreatment than children who had not previously been maltreated (Chi-sq=19.4, df=2, p<0.01); authors' estimated odds ratio=5.96, follow-up at two years. Recurrence of neglect is also associated with prior history of abuse/neglect (Chi-sq=13.6, df=2, p=0.01) (Wood 1997, study quality score 10). 9. Number of prior abuse episodes was found not to be significantly associated with re-abuse in two related studies (DePanfilis and Zuravin 1999a, 2002, no statistical data reported). <u>Risk factor - child factors</u> Child factors were examined in 8 studies (English et al. 1999; Fluke et al. 1999; Fryer and Miyoshi 1994; Herrenkohl et al. 1979; Murphy et al. 1992; Rittner 2002; Rivara 1985; Swanston et al. 2002). <u>Risk factor - child factors – age</u> Seven studies looked at the impact of child age (English et al. 1999; Fluke et al. 1999; Fryer and Miyoshi 1994; Herrenkohl et al. 1979; Murphy et al. 1992; Rivara 1985; Swanston et al. 2002). The review authors report that 4 found that younger children were at higher risk (English et al. 1999; Fluke et al. 1999; Fryer and Miyoshi 1994; Herrenkohl et al. 1979). However, in Table 1, 1 of these is reported as non-significant (Fluke et al. 1999). Three found no association with age (Murphy et al. 1992; Rivara 1985; Swanston et al. 2002). 1. Younger children were at higher risk of recurrence of maltreatment (p<0.05), follow-up at 18 months (English 1999, study quality score 9).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		2. No significant association between recurrence and age. Review does not report statistical data. (Fluke et al. 1999). 3. Younger children more vulnerable, follow-up at 4 years; p<0.001. (Fryer 1994, study quality score 8). Higher rates of recurrence in families with a child aged 0–5 years and lower rates with children over 11 (Chi-sq=23.37, df=2, p<0.01), follow-up at 10 years (Herrenkohl 1979, poor study quality score 4). 5. Children's age was not associated with return to court for abuse offence, follow-up at 3.5 years. (No statistical data reported) (Murphy 1992, study quality score 4). 6. There was no relationship between recurrent maltreatment and the child's age, follow-up at 31 months. (No statistical data reported). (Rivara 1985, study quality score 6). 7. Demographic factors (including age) not significantly related to recurrence, follow-up at 6 years. (Swanston 2002, study quality score 9, no statistical data reported). <u>Risk factor - child factors – gender/sex</u> Gender differences were examined in 3 studies (Fryer and Miyoshi 1994; Rittner 2002; Swanston 2002). Based on data in text, the author states that no significant association was found between sex of children and abuse recurrence. 1. No significant differences in recurrence between boys and girls, follow-up at 4 years, 18 months and 6 years respectively. (Fryer and Miyoshi 1994, study quality score 8; Rittner 2002, study quality score 10; Swanston 2002, study quality score 9, no statistical data reported).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Risk factor - child factors – ethnicity One study examined the relationship between ethnicity and abuse recurrence (Fryer 1994), and found no significant association. 1. No significant association was reported between ethnicity/race and maltreatment recurrence, follow-up at 4 years. (Fryer 1994, study quality score 8, no statistical data reported). Risk factor - parental factors – caregivers' abuse history Four studies examined the association between recurrence of maltreatment and the child's primary caretaker themselves having been maltreated as a child (English 1999; Rittner 2002; Swanston 2002; Wood 1997). A positive association was reported in three studies (English 1999; Rittner 2002; Wood 1997) but such an association was not found in one study (Swanston 2002, based on data presented in Table 1.) 1. A significant association between higher rate of recurrence of maltreatment and primary caregiver abused as a child ($p<0.05$), follow-up at 18 months (English 1999, study quality score 9). 2. A significant association between higher rate of recurrence of maltreatment and caretaker's own abuse history, especially neglect (Chi-sq=11.08, df=1, $p<0.001$), follow-up at 18 months (Rittner 2002, study quality score 10). 3. A significant association between higher rate of recurrence of maltreatment and primary caregiver abused as a child (Chi-sq=6.0, df=1, $p=0.01$), follow-up at 2 years (Wood 1997, study quality score 10).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		4. No significant relationship between recurrence of sexual abuse and caregiver's abuse history as a child, follow-up at 6 years (no statistical data reported) (Swanston 2002, study quality score 9). <u>Risk factor - parental factors - caregiver's substance abuse</u> Three studies examined the association between a parental history of substance abuse and maltreatment recurrence, based on data presented in text and in table 1 (English 1999; Rittner 2002; Swanston 2002). A significant association was found between these two relationships in these 3 studies. 1. A significant association between higher rate of recurrence of maltreatment and parental history of substance abuse ($p<0.05$), follow-up at 18 months (English 1999, study quality score 9). 2. A significant association between higher rate of recurrence of maltreatment and parental history of alcohol abuse (risk ratio=2.67, 95% CI 1.24–5.74), follow-up at 6 years (Swanston 2002, study quality score 9). 3. A significant association between a parental history of alcohol abuse and subsequent maltreatment, follow-up at 18 months. This association disappeared when other factors were controlled for using multivariate analysis. (Based on author's report in text, no data was reported in table 1, Rittner 2002, study quality score 10). <u>Risk factor - parental factors - caregiver's mental health problems</u> Four studies examined the association between parental mental problems and maltreatment recurrence	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		(Rittner 2002; English 1999; Murphy 1992; Swanston 2002). The author concluded that parental mental health problems is consistently identified as a factor predicting future maltreatment. 1. An association was found between higher rate of reabuse and parental mental health problems in multivariate analysis (no statistical data reported), follow-up at 18 months (Rittner 2002, study quality score 10). 2. A significant association was found between higher rate of recurrence of maltreatment and parental mental health problems (risk ratio=4.23, 95% CI 2.01–8.89), follow-up at 6 years (Swanston 2002, study quality score 9). 3. A significant association ($p<0.05$) was found between higher rate of recurrence of maltreatment and primary caregiver impairments (mental, physical, emotional — not further specified), follow-up at 18 months (English 1999, study quality score 9). 4. A significant association was found between higher rate of recurrence of maltreatment and primary caregiver mental health problems (psychosis, character disorders) (Chi-sq=5.4, df=1, $p<0.05$), follow-up at 3.5 years (Murphy 1992, poor study quality score 4). <u>Risk factor - parental factors - Primary caregiver intellectual limits</u> One study examined the association between primary caregivers' intellectual limits and subsequent neglect (Wood 1997). This study found a significant association between the two relationships.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		1. A significant association was reported between higher rate of subsequent neglect and primary caregiver intellectual limits (Ci-sq = 8.8, df=1, p = 0.01), follow-up at 2 years, but caretaker age was not associated with recurrent maltreatment. (Wood 1997, study quality score 10). <u>Risk factor - parental factors - Parenting ability/skills</u> One study examined the relationship between parenting ability and recurrent maltreatment (Johnson and L'Esperance 1984), and this study found a significant association between the two relationships. The authors concluded that parental conflict is a factor consistently identified as predicting future maltreatment. 1. A significant association (r=0.36, p<0.0005) was reported between higher rate of recurrence of maltreatment and parenting skills (mothering skills), follow-up at 2 years and based on data in Table 1. (Johnson and L'Esperance 1984, study quality score 9). <u>Risk factor - parental factors - caregiver support</u> One study examined the relationship between caregiver support and recurrent maltreatment (English 1999) and found that the degree of protection offered to the child by the non-abusing carer was linked to lower rates of recurrent maltreatment. 1. The degree of protection offered to the child by the non-abusing carer was linked to lower rates of recurrent maltreatment, follow-up at 18 months. Based on data reported in text. (English 1999, study quality score 9, no statistical data reported).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Risk factor - family environmental factors - parental conflict Four studies examined the relationship between parental conflict and maltreatment recurrence, based on data presented in Tables 1 and 2 (DePanfilis and Zuravin 1999a, 2002; English 1999; Swanston 2002). NB. Please note discrepancies as the 3 studies quoted in the text did not match the 4 studies presented in Tables 1 and 2). The authors conclude that parental conflict is a factor consistently identified as predicting future maltreatment. 1. A significant association was reported between higher rate of recurrence of maltreatment and parental conflict (risk ratio=2.25, 95% CI 1.1–4.62), follow-up at six 6 (Swanston 2002, study quality score 9). 2. A significant association was found between higher rate of recurrence of maltreatment and domestic violence ($p<0.05$), follow-up at 18 months (English 1999, study quality score 9). 3/4. A significant association was found in two related studies between higher risk of recurrent maltreatment and families who had a child previously placed in care (risk ratio=1.9, $p=0.002$), follow-up at 5 years (DePanfilis and Zuravin 1999a, 2002, study quality score 12). Risk factor - family environmental factors - Change in caregiver before intake One study examined the association between a change in caregiver before intake and maltreatment recurrence (Swanston 2002) and found a significant association between the two relationships.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		1. A significant association between higher rate of recurrence of sexual abuse and maltreatment and a change in caregiver before intake (Chi-sq = 17.77; df = 2; p = 0.001), social workers' rating of family functioning (Chi-sq = 11.27, df = 4, p = 0.02), and multiple changes in caregiver (Chi-sq= 17.44, df = 1, p<0.001), follow-up at six years (Swanston 2002, study quality score 9). <u>Risk factor - family environmental factors - number of victims involved in an incident of abuse</u> One study examined the relationship between the number of victims involved in an incident of abuse and abuse recurrence (Wood 1997), and found a significant association between the two variables. 1. Number of victims involved in an incident of abuse was significantly associated with recurrent maltreatment (Chi-sq=8.8, df=1, p=0.01); authors' estimated odds ratio = 5.96, follow-up at 2 years (Wood 1997, study quality score 10). <u>Risk factor - family environmental factors - support and supervision</u> Three studies examined the association between inadequate supervision and abuse recurrence (De-Panfilis and Zuravin 1999a; English 1999; Wood 1997). The authors conclude that a deficit in social support is significantly associated with recurrent maltreatment. 1. Inadequate supervision by either caregiver was significantly associated with subsequent maltreatment (Chi-sq=4.6, df=1, p=0.03); follow-up at 2	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		years. The same study found no significant association between recurrent maltreatment and inadequate physical (Wood 1997, study quality score 10, no statistical data reported.) 2. Significant associations (risk ratio=1.4, p=0.0001) were reported between higher rate of recurrence of maltreatment and social support deficit construct (no support in extended family, no supportive friends, ineffective use of informal helping systems), between recurrence of maltreatment and family stress construct (risk ratio=1.2, p=0.02), between recurrence of maltreatment and child vulnerability construct (risk ratio=1.4, p = 0.02), follow-up at 5 years (DePanfilis and Zuravin 1999a, study quality score 12). 3. A significant association was found between higher risk of recurrent maltreatment and lack of social support (p<0.05), follow-up at 18 months (English 1999, study quality score 9). <u>Risk factor - family environmental factors – parental stress</u> One study examined the relationship between parental stress and abuse recurrence (Johnson and L'Esperance 1984) and found a significant association between the 2 variables. 1. A significant association was found between higher risk of recurrent maltreatment and parental stress (>1 child in home) (r=0.26, p<0.001), follow-up at 2 years (Johnson and L'Esperance 1984, study quality score 9). <u>Risk factor - family environmental factors – economic factors</u>	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		One study examined the relationship between economic factors and abuse recurrence (Rittner 2002) and found a significant association between the two variables. 1. A significant association was reported between higher rate of recurrence of maltreatment and having no income, follow-up at 18 months. (Rittner 2002, study quality score 10, no statistical data reported). <u>Risk factor - engagement with services</u> Five studies examined the relationship between service engagement and maltreatment recurrence (DePanfilis and Zuravin 2002; Johnson and L'Esperance 1984; Littel 2001; Rittner 2002; Rivara 1985). A positive association was found in 3 studies (DePanfilis and Zuravin 2002; Johnson and L'Esperance 1984; Littel 2001). No such association was found in 2 studies (Rittner 2002; Rivara 1985). One study did not find an association between reduced recurrence and admission by perpetrator; numbers of caseworkers or casework contacts; use of the juvenile court; level of cooperation of caregiver; presence of signed service agreement (DePanfilis and Zuravin 2002). 1. A significant association (risk ratio=0.688, p=0.05) was reported between reduced risk of recurrence of maltreatment and attendance at Child Protective Services ('… attendance reduces risk of recurrence by 32% …' (p746), follow-up at 5 years. The same study found no significant relationship between reduced risk of recurrence of maltreatment and admission by perpetrator; numbers of caseworkers or casework contacts; use of the juvenile court; level of	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		cooperation of caregiver; presence of signed service agreement; and degree of improvement by the end of the study. (DePanfilis and Zuravin 2002, study quality score 12, no statistical data reported). 2. A significant association (r=0.33, p<0.0005) was reported between reduced risk of recurrence of maltreatment and a '… client's capacity to use resources …' (p747), follow-up at 2 years (Johnson and L'Esperance 1984, study quality score 9). 3. A direct relation was found between participation in treatment planning ('collaboration') and compliance with programme expectations; compliance (keeping appointments, completing tasks, and cooperation) was associated with a small reduction in substantiated report during Family Preservation Services (b=-0.08), but not after Family Preservation Services; follow-up at one year (Littel 2001, study quality score 11). 1 4. No association was found between recurrence of maltreatment and cooperation/compliance with court orders, follow-up at 18 months (Rittner 2002, study quality score 10, no statistical data reported). 5. No association between recurrence of maltreatment and compliance with treatment, follow-up at 30.8 months (Rivara 1985, study quality score 6, no statistical data reported).	

9. Jones L, Bellis MA, Wood S et al. (2012) Prevalence and risk of violence against children with disabilities: A systematic review and meta-analysis of observational studies. Lancet 380: 899–907

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To '… synthesise evidence for the prevalence and risk of violence against	**Participants:** Children and young people. Children with a range of disabilities and non-disabled comparison children. No further details reported.	Although this appears to be a well conducted review and meta-analysis it should be noted that very little detail is provided in relation to the characteristics of participants in the individual studies, and in some	**Overall assessment of internal validity:** ++ Although this appears

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
children with disabilities' (p899) The review reports on a meta-analysis of studies in which the prevalence or risk of violence was reported in children with disabilities only, and those in which prevalence or risk in disabled children was compared to prevalence or risk in non-disabled children. As the NCCSC has focused on comparative studies in relation to questions on recognition only data from studies with a non-disabled comparison group has been reported by the NCCSC. **Methodology:** Systematic review. Systematic review and meta-analysis of cross-sectional, case-control, or cohort studies. Included studies relevant to question 3 of the NCCSC review had either a cross-sectional or a cohort design.	**Sample characteristics:** • Age - Not reported for any of the included studies applicable to question 3 of the NCCSC review. • Sex - Not reported for any of the included studies applicable to question 3 of the NCCSC review. • Ethnicity - Not reported for any of the included studies applicable to question 3 of the NCCSC review. • Religion/belief - Not reported for any of the included studies applicable to question 3 of the NCCSC review. • Disability - Disabilities included physical and sensory impairments, and mental illness. • Long term health condition - Not reported for any of the included studies applicable to question 3 of the NCCSC review. • Sexual orientation - Not reported for any of the included studies applicable to question 3 of the NCCSC review. • Socioeconomic position - Not reported for any of the included studies applicable to question 3 of the NCCSC review. • Type of abuse - Studies measured physical violence, sexual violence, emotional abuse, neglect, or combinations of any of these.	cases the information that is provided suggest that the definition of abuse used by each study may be quite wide. Findings should therefore be used with caution. NB Only data reported from comparative studies has been extracted by the NCCSC. <u>Association between any disability and risk of any type of violence</u> Spencer et al. (2005) (autism): There was a non-significant association between autism and any type of violence, however this was non-significant; odds ratio=0.82 (95% CI 0.30–2.19). Verdugo et al. (1995): Disabled children were found to be at significantly increased risk of any type of violence; odds ratio=8.56 (95% 3.61–24.66). Spencer et al. (2005) (vision or hearing): There was a non-significant association between vision or hearing impairments and any type of violence; odds ratio=0.87 (95% CI 0.36–2.11). Spencer et al. (2005) (psychological problems): Children with psychological problems were found to be at significantly increased risk of any type of violence; odds ratio=5.24 (95% CI 2.14–8.74). Spencer et al, (2005) (cerebral palsy): Children with cerebral palsy were found to be at significantly increased risk of any type of violence; odds ratio=3.12 (95% CI 1.70–5.72). Cuevas et al. (2009): Disabled children were found to be at significantly increased risk of any type of violence; odds ratio=1.75 (95% 1.23–2.45).	to be a well-conducted review and meta-analysis it should be noted that very little detail is provided in relation to the characteristics of participants in the individual studies, and in some cases the information that is provided suggest that the definition of abuse used by each study may be quite wide. Findings should therefore be used with caution. **Overall assessment of external validity:** ++ **Overall validity rating:** ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Country: Not reported. It is not possible to determine the countries which included studies (with a non-disabled comparison group) were conducted in. **Source of funding:** Other – WHO Department of Violence and Injury Prevention and Disability.	• Looked after or adopted status - Not reported for any of the included studies applicable to question 3 of the NCCSC review. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported for any of the included studies applicable to question 3 of the NCCSC review. **Recognition indicators measured:** Risk factors - Disabilities (e.g. physical impairments, mental illness, sensory impairments).	Spencer et al. (2005) (behaviour disorder): Children with behaviour disorders were found to be at significantly increased risk of any type of violence; odds ratio=11.48 (95% CI 8.52–15.46). Spencer et al. (2005) (speech or language): Children with speech or language disorders were found to be at significantly increased risk of any type of violence; odds ratio=3.26 (95% 2.44–4.34). Spencer et al. (2005) (learning difficulties): Children with learning disabilities were found to be at significantly increased risk of any type of violence; odds ratio=6.50 (95% 5.25–8.09). Sullivan et al. (2000): Disabled children were found to be at significantly increased risk of any type of violence; odds ratio=4.53 (95% 4.17–4.93). Overall (random effects pooled odds ratios; I^2=91.8%): Children with any type of disability were found to be at significantly increased risk of any type of maltreatment; odds ratio=3.68 (95% 2.56–5.29). NB The authors refer to both maltreatment and violence in relation to this data. See below for further details. <u>Association between any disability and risk of physical violence</u> Spencer et al. (2005) (vision or hearing): There was a non-significant association between vision or hearing impairments and physical violence; odds ratio=0.52 (95% 0.07–3.73).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Spencer et al. (2005) (autism): Children with autism were found to be at increased risk of physical violence, however this was non-significant; odds ratio=1.23 (95% 0.31–4.96). Spencer et al. (2005) (psychological problems): Children with psychological problems were found to be at significantly increased risk of physical violence; odds ratio=3.75 (95% 1.39–10.12). Reiter et al. (2007): Disabled children were found to be at increased risk of physical violence, however this was non-significant; odds ratio=1.30 (95% 0.53–3.23). Spencer et al. (2005) (cerebral palsy): Children with cerebral palsy were found to be at significantly increased risk of physical violence; odds ratio=5.08 (95% 2.25–11.47). Miller (1996): Disabled children were found to be at significantly increased risk of physical violence; odds ratio=3.05 (95% 1.49–6.26). Dawkins (1996): Disabled children were found to be at increased risk of physical violence, however this was non-significant; odds ratio=2.67 (95% 0.81–3.23). Cuevas et al. (2009): Disabled children were found to be at significantly increased risk of physical violence; odds ratio=2.46 (95% 1.30–4.45).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Spencer et al. (2005) (behaviour disorder): Children with behaviour disorders were found to be at significantly increased risk of physical violence; odds ratio=6.44 (95% 3.52–11.80). Spencer et al. (2005) (learning difficulties): Children with learning difficulties were found to be at significantly increased risk of physical violence; odds ratio=3.87 (95% 2.47–6.07). Spencer et al. (2005) (speech or language): Children with speech or language disorders were found to be at significantly increased risk of physical violence; odds ratio=4.92 (95% 3.28–7.38). Sullivan et al. (2000): Disabled children were found to be at significantly increased risk of physical violence; odds ratio=4.35 (95% 3.88–4.86). Overall (random effects pooled odds ratios; I^2 =50.6%): Children with any type of disability were found to be at significantly increased risk of physical violence; odds ratio=3.56 (95% 2.80–4.52). See below for further details. The authors report that the exclusion of two outliers (Reiter et al. 2007, and data relating to children with vision or hearing impairments reported by Spencer et al. 2005) resulted in a larger pooled odds ratio of 4.05 (95% CI 3.39–4.82). It is not clear why these data were considered to be outliers. <u>Association between any disability and risk of sexual violence</u> Spencer et al. (2005) (psychological problems): Children with psychological problems were found to be	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		at increased risk of sexual violence, however this was non-significant; odds ratio=2.32 (95% 0.32–16.57). Spencer et al. (2005) (speech or language): Children with speech or language disorders were found to be at increased risk of sexual violence, however this was non-significant; odds ratio=1.40 (95% 0.45–4.39). Reiter et al. (2007): Disabled children were found to be at significantly increased risk of sexual violence; odds ratio=3.50 (95% 1.25–10.36). Miller (1996): Disabled children were found to be at significantly increased risk of sexual violence; odds ratio=7.30 (95% 3.11–18.03). Spencer et al. (2005) (behaviour disorder): Children with behaviour disorders were found to be at significantly increased risk of sexual violence; odds ratio=10.27 (95% 4.81–21.94). Spencer et al. (2005) (learning difficulties): Children with learning difficulties were found to be at significantly increased risk of sexual violence; odds ratio=8.03 (95% 4.82–13.38). Cuevas et al. (2009): Disabled children were found to be at increased risk of sexual violence, however this was non-significant; odds ratio=1.51 (95% 0.94–2.35).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Suris et al. (1996): Disabled children were found to be at significantly increased risk of sexual violence; odds ratio=1.78 (95% 1.43–2.23). Alriksson-Schmidt et al. (2010): Disabled children were found to be at significantly increased risk of sexual violence; odds ratio=2.35 (95% 1.94–2.83). Everett Jones et al. (2008): Disabled children were found to be at significantly increased risk of sexual violence; odds ratio=2.64 (95% 2.24–3.11). Blum et al. (2001): Disabled children were found to be at significantly increased risk of sexual violence; odds ratio=1.87 (95% 1.60–2.19). Sullivan et al. (2000): Disabled children were found to be at significantly increased risk of sexual violence; odds ratio=3.31 (95% 2.87–3.79). Overall (random effects pooled odds ratios; I^2=86.9%): Children with any type of disability were found to be at significantly increased risk of sexual violence; odds ratio=2.88 (95% 2.24–3.69). See below for further details. <u>Random-effects pooled odds ratios for risk of violence (odds ratios and 95% confidence intervals)</u> I^2 statistic (95% CI) used to estimate heterogeneity between pooled studies. Association between any disability and risk of any type of maltreatment: Children with any type of disability were found to be at significantly increased risk	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		of any type of violence; data pooled from 4 studies; odds ratio=3.68 (95% 2.56–5.29); heterogeneity 91.8% (95% 87.7–94.1). NB The authors refer to both maltreatment and violence in relation to this data. Association between any disability and risk of physical violence: Children with any type of disability were found to be at significantly increased risk of physical violence; data pooled from 6 studies; odds ratio=3.56 (95% 2.80–4.52); heterogeneity 50.6% (95% 0-73.0). However, bias assessment showed asymmetry in the funnel plot (Egger test, p=0.01; Begg-Mazumdar test, p=0.04) and the authors report that the exclusion of two outliers (Reiter et al. 2007), and data relating to children with vision or hearing impairments reported by Spencer et al. (2005) resulted in a larger pooled odds ratio of 4.05 which was also significant (95% CI 3.39–4.82). It is not clear why these data were considered to be outliers. Association between any disability and risk of sexual violence: Children with any type of disability were found to be at significantly increased risk of sexual violence; data pooled from 9 studies; odds ratio of 2.88 (95% 2.24–3.69); heterogeneity 86.9% (95% 78.8–90.9). Association between any disability and risk of emotional abuse: Children with any type of disability were found to be at significantly increased risk of emotional abuse; data pooled from 4 studies; odds ratio=4.36 (95% 2.42-7.87); heterogeneity 94.4 (95% 91.4-96.0).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Association between any disability and risk of neglect: Children with any type of disability were found to be at significantly increased risk of neglect; data pooled from 3 studies; odds ratio = 4.56 (95% 3.23-6.43); heterogeneity 73.8% (95% 27.7-86.0). Association between mental or intellectual disability and any type of maltreatment: Children with a mental or intellectual disability were found to be at significantly increased risk of any type of maltreatment; data pooled from 3 studies; odds ratio=4.28 (95% 2.12–8.62); heterogeneity 94.0% (95% 90.2–95.9). NB The authors refer to both any maltreatment and any violence in relation to this data. Association between mental or intellectual disability and physical violence: Children with a mental or intellectual disability were found to be at significantly increased risk of physical violence; data pooled from 4 studies; odds ratio=3.08 (95% 2.08–4.57); heterogeneity 50.8% (95% 0–77.2). Association between mental or intellectual disability and sexual violence: Children with a mental or intellectual disability were found to be at significantly increased risk of sexual violence; data pooled from 4 studies; 4.62 (95% 2.08–10.23); heterogeneity 84.7% (95% 64.4–91.2). Association between mental or intellectual disability and emotional abuse: Children with a mental or intellectual disability were found to be at significantly increased risk of emotional abuse; data pooled from	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		3 studies; odds ratio=4.31 (95% 1.37–13.56); heterogeneity 96.2% (95% 94.2-97.3). Association between mental or intellectual disability and neglect: Data pooled from 2 studies did not produce a sample size large enough to calculate pooled odds ratios. Potential sources of heterogeneity: The authors report that visual inspection of the forest plot suggested sample size as a potential source of heterogeneity; however univariate meta-regression analyses showed that sample size (as a continuous covariate) did not have a significant impact on risk of violence estimates (statistical data not presented). For estimates of risk of physical violence, analysis showed that the method of reporting (official records vs. self-report) had a significant impact on risk estimates, 0.60 vs. 0.21, p=0.02. For estimates of risk of sexual violence, analysis showed that the type of disability (mental or intellectual disability vs. other types of disability) had a significant impact on risk estimates, 0.76 vs. 0.33, p=0.05).	

10. Kočovská E, Puckering C, Follan M et al. (2012) Neurodevelopmental problems in maltreated children referred with indiscriminate friendliness. Research in Developmental Disabilities 33: 1560–5

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The study aimed to '… explore the extent of neurodevelopmental difficulties in severely maltreated adopted children'	**Participants:** Children and young people. The authors aimed to compare a group of adopted children with experience of severe maltreatment early in their life (now living in a stable environment) with symptoms of	<u>Intelligence (verbal and performance tested using the Wechsler Abbreviated Scale of Intelligence. Exact scales used unclear)</u> <u>Verbal IQ</u>: Children in the adopted group had significantly lower scores than children in the comparison group on measures of verbal IQ; adopted group t=-	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
(p1560). We have extracted only data relating to language. **Methodology:** Cross-sectional study. **Country:** UK – Scotland – Glasgow. **Source of funding:** • Government – NHS Greater Glasgow and Clyde. • Other – University of Gothenburg.	indiscriminately friendly behaviour, to a group of non-maltreated '… typically developing children… ' (p1561). The inclusion criteria for the adopted group were '… symptoms of indiscriminately friendly behaviour plus a history of maltreatment' (p1561). Exclusion criteria for the adopted group were moderate or severe intellectual disability (which can also lead to disinhibited behaviour), and current experience of maltreatment or family instability. The adopted group was recruited through an adoption charity, which approached eligible families living within travelling distance of the clinic. The authors note that only a small proportion of children in this group had had any contact with Child and Adolescent Mental Health Services and of those currently accessing this service, the majority were not in receipt of ongoing therapy. The only inclusion criterion for the comparison group was aged between 5 and 12 years. Exclusion criteria for the comparison group were – any psychiatric diagnosis, a history of maltreatment (including suspected), known involvement	3.41; p=.001. The reviewing team calculated effect size using reported data; ES=-1.14. <u>Verbal-performance</u>: The authors reports that children in the adopted group had significantly lower scores than children in the comparison group on measures of verbal-performance IQ; t=0.73; p=.001. However, the t value reported here appears to be in error – calculation by the reviewing team using reported means and standard deviation found t=0.153, p=0.88. The reviewing team calculated effect size using reported data; ES=-0.04. <u>Language ability, narrative speech, and short term-memory (tested using the Renfrew Language Scales – Bus Story Test)</u> - Performance below chronological age - The number of children in the adopted group performing below their chronological age on measures of language, speech and short-term memory was significantly higher than the number in the comparison group (chi-square=not reported, p=.001). It was not possible to calculate effect sizes from the reported data. The authors also report narratively that 'Over half of the adopted children had suspected language disorder and/or delay, on the Renfrew Bus Test, in comparison to 10% of the comparison group children' (p1564) Need for full assessment: The number of children in the adopted group whose language difficulties were deemed to 'merit' full assessment was significantly higher than the number in the comparison group;	**Overall validity rating:** + Little justification given for choice of statistical tests.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	with 'social work', child protection registration, or trauma within the past year. The comparison group was recruited via letter through two general practice surgeries that were determined to have 615 potentially eligible children registered (aged between 5 and 12). Due to initial imbalances in relation to age and gender, a second round of recruitment letters were sent to families of boys between the ages of 6 and 10 who had not responded to the first recruitment letter. **Sample characteristics:** • Age - All children were between the ages of 5 and 12 years. Adopted group - mean age = 9.4 years (1.8 SD). Comparison group – mean age = 8.7 (2.4). • Sex - Adopted group 51.5% male (n=18). Comparison group 43.1% male (n=17). • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Pre-adoption histories (i.e. in relation to history of maltreatment, birthweight, etc.) of the	adopted group (Fisher's exact test, value not reported; p=.002). It was not possible to calculate effect sizes from the reported data.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	adopted group was extracted from social worker notes using a check-list designed specifically for the study. The adopted group had experience of the following before their adoption: Alcohol misuse by a birth parent = 74%. Drug misuse by a birth parent = 62%. Emotional and/or physical neglect by a birth parent = 100%. History of physical abuse = 49%. History of sexual abuse = 20%. • Looked after or adopted status - Mean age at adoption for the adopted group was 62.9 months (25.3 SD); mean number of months with adoptive family was 51.3 (26.8 SD). • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Adopted group – Referred n=43; assessed as eligible n=39; clinically assessed n=34 (2 families/5 children withdrew). Although the authors state that 34 children were clinically assessed data only appears to be presented for 33. • Comparison group – A total of 32 children were clinically assessed. 461 recruitment letters sent to eligible families/children; responses re-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	ceived n=58; withdrew n=9; clinically assessed – number unclear (authors report that not all of the remaining children were assessed due to gender and age mismatches). To address these another 62 recruitment letters were re-sent to families who had not originally responded (only those with male children between the ages of 6 and 10). Four had moved, and 6 responded and were clinically assessed. Sample size – The total numbers of children assessed was n=66. **Recognition indicators measured:** • Language - Intelligence (verbal, performance and full) was tested using the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999). • Language - Language ability, narrative speech, and short term-memory were tested using the Renfrew Language Scales – Bus Story Test (Renfrew 1991).		

11. Lereya T et al. (2013) Parenting behavior and the risk of becoming a victim and a bully/victim: a meta-analysis study. Child Abuse & Neglect 37: 1091–108

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: 'The objective of this meta-analysis is to systematically investigate the type and strength of the association between parenting behaviour ... on being bullied' (p1092). **Methodology:** Systematic review of a total of 70 studies. Six studies had specific relevance to abuse and neglect, and were analysed separately. Only data relating to these six studies are reported here. **Country:** Range of countries. Europe (4 studies, no further detail on specific countries), US (1 study), Other (1 study). **Source of funding:** Other – Economic and Social Research Council and Qatar National Research Fund.	**Participants:** Children and young people - Studies included in the meta-analysis involved children between the ages of 4 and '12+'. The authors do not specify what the upper bound for 12+ is for these studies. **Sample characteristics:** • Age - Bowes et al. (2009) 4 to 7 Dehue et al. (2012) 7.5 to 12 Kelleher et al. (2008) 12 upwards Mohr (2006) 12 upwards Schwartz et al. (2000) (studies 1 and 2) 7.5 to 12 Shin and Kim (2008) 12 upwards. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported.	**Statistical data -** 1. Association between abuse and neglect and being a victim of bullying. This was explored in 6 studies, with the following effect sizes, as calculated using Hedge's g, and 95% confidence intervals. Bowes et al. (2009) - Hedge's g=0.444 (95% CI 0.247 to 0.641) Dehue et al. (2012) - Hedge's g=0.195 (95% CI 0.041 to 0.350) Kelleher et al. (2008) - Hedge's g=0.097 (95% CI -0.538 to 0.732) Mohr (2006) - Hedge's g=0.555 (95% CI 0.104 to 1.006) Schwartz et al. (2000) - Hedge's g=0.386 (95% CI 0.169 to 0.604) Shin and Kim (2008) - Hedge's g=0.081 (95% CI -0.267 to 0.429). Two of the 4 studies had 95% confidence intervals which crossed the zero threshold, indicating a non-significant result. The combined effect size of the 6 studies was Hedge's g=0.307 (95% CI 0.175 to 0.440). This suggests that, overall, children who had experienced abuse and neglect were more likely to be the victims of bullying. Publication bias: Failsafe n=42 and exceeded Rosenthal's 5k+10 benchmark=40, suggesting low risk of publication bias. 2. Association between abuse and neglect and being a bully/victim This was explored in 3 studies, with the following effect sizes, as calculated using Hedge's g, and 95% confidence intervals. Bowes et al. (2008) - Hedge's g=0.748 (95% CI 0.520 to 0.976) Dehue et al. (2012) - Hedge's g=0.440 (95% CI 0.054 to 0.827) Mohr (2006) - Hedge's g=1.010 (95% CI 0.440 to 0.919) No studies had 95% confidence intervals which crossed the zero threshold. The combined effect size of the three studies was Hedge's g=0.680 (95% CI 0.440 to 0.919). This suggests that, overall, children who had experienced abuse and neglect	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + Key limitations: No critical appraisal of included studies. However, the rest of the systematic review is of high quality, and statistical data is well reported. This has therefore been rated as 'moderate' quality.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Sample size: • Sample size - Total sample size across six studies=5,289. • Systematic reviews: number of studies. Six studies were relevant to our review question. **Recognition indicators measured:** • Bullying. Study examines association between abuse and neglect and being a victim of bullying and being a 'bully/victim' - individuals who both bully others and are victims of bullying (e.g. Wolke and Samara 2004).	were more likely to be bully/victims than children who had not experienced abuse and neglect. Publication bias: Failsafe n=30 and exceeded Rosenthal's 5k+10 benchmark=25, suggesting low risk of publication bias. **Narrative findings** A meta-analysis of 6 studies comparing rates of experiencing bullying in children (total n=5289, age ranges from 4 to 12+) who had been abused or neglected compared to those who had not, found that children who had been abused or neglected were more likely to be the victims of bullying, with small effect size (Hedge's g=0.307 (95% CI 0.175 to 0.440)). A meta-analysis of 3 studies comparing rates of being a bully/victim in children (total n=4149, age ranges from 4 to 12) who had been abused or neglected, compared to those who had not, found that children who had been abused or neglected were more likely to be bully/victims, with medium to large effect size (g=0.680 (95% CI 0.440 to 0.919)).	

12. Luke N, Banerjee R (2013) Differentiated associations between childhood maltreatment experiences and social understanding: A meta-analysis and systematic review. Developmental Review 33: 1–28

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To '… evaluate the strength of evidence for the hypothesis that physically abused or neglected children underperform relative to their nonmaltreated peers in	**Participants:** Children and young people - Maltreated children. **Sample characteristics:** • Age - The review protocol does not state that the age of the samples used in the individual studies was used as a screening criterion. The authors classified the	**Statistical data (negative effect size corresponds to poorer performance by maltreated participants)** Meta-analysis - overall (negative effect size corresponds to poorer performance by maltreated participants) – 16 of the 19 studies (84.2%) showed effect sizes in the expected direction, that is maltreatment status or severity was associated with poorer emotion skills although only 12 of these were significant.	**Overall assessment of internal validity:** + The inclusion of studies in which the sample was partially or wholly comprised of adults (and the lack of discussion in relation to this decision),

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
measures of social understanding' (p2). **Methodology:** Systematic review and meta-analysis. NB. The NCCSC have only extracted data from the meta-analysis (random effects model). Studies included in the meta-analysis compared groups of maltreated participants with groups of non-maltreated participants or used prevalence or severity of maltreatment as a continuous variable. The wider review does not include quantitative data and the findings of this have therefore not been extracted. **Country:** Range of countries. The studies included in the meta-analysis were conducted in: Barahal et al., 1981: USA. Bowen and Nowicki, 2007: United Kingdom. Camras et al., 1983: USA. Camras et al., 1988: USA. During	studies by age into 3 age ranges in order to enable moderator analysis. These were - early childhood (ages 2–6); middle childhood (ages 7–11); and adolescence and adulthood (ages 12 and over). Samples in which age ranges overlapped were classified according to the category in which most of the age range fell. Where the age range was spread equally between categories, classification was based on the sample's mean age. The total age range for studies included in the meta-analysis was 2 years and 8 months to 74 years. The age range for studies which only included participants under 18 was 2 years and eight months to 18 years. The age range for studies which included adults was 18 to 74 years (this is based on one study – Gapen 2010, and it is not clear why this study was included given that the sample was comprised wholly of adults. The authors classify the age range of this sample as 'adolescence and adulthood'. The ages of the samples for each individual study used in the meta-analysis were - Barahal et al. 1981: 6–8 years	Three of the 19 studies showed effect sizes in the reverse direction, that is maltreatment status or severity was associated with better emotion skills, although only one of these was significant. The overall mean effect size across the 19 studies showed a medium effect size in the direction of maltreated children showing poorer emotion skills: d=-0.696; SE=.148; 95% CI -0.985 to -0.406; Z=-4.714; p<.001; Q=131.331 (between studies); df (Q) 18; p<.001. Meta-analysis – moderated by outcome variable (negative effect size corresponds to poorer performance by maltreated participants) – A moderator analysis examining the effect of choice of outcome variable was conducted. The results suggested that the type of outcome measure did moderate the findings (Q(2)=13.001, p=0.002), with studies measuring emotion understanding showing larger effect sizes than those measuring composite emotion knowledge, which in turn were larger than those measuring emotion recognition: Emotion understanding: d=-1.351; 95% CI -2.311 to -0.392; Z=-2.760; p=.006. Emotion knowledge (composite of emotion recognition and emotion understanding): d=-0.972; 95% CI -1.258 to -0.686; Z =-6.660; p<.001. Emotion recognition: d=-0.309; 95% CI -0.580 to -0.039; Z=-2.239; p=.025. The authors note that this may be because emotion understanding is a more advanced skill, and so may be 'particularly susceptible to the deleterious effects of maltreatment experiences' (p20). Post-hoc comparisons showed that the difference between studies measuring emotion understanding and	the lack of information on the quality of included studies, and the small number of databases searched are areas of concern. **Overall assessment of external validity:** ++ **Overall validity rating:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
and McMahon, 1991: USA. Edwards et al., 2005: USA. Gapen, 2010: USA. Leist and Dadds, 2009: Australia. Pajer et al., 2010: USA. Pears and Fisher, 2005: USA. Perlman, Kalish, and Pollak, 2008: USA. Pollak et al., 1997: USA. Pollak et al., 2001: USA. Shackman and Pollak, 2005: USA. Shipman and Zeman, 1999: USA. Shipman et al., 2005: USA. Smith and Walden, 1999: USA. Sullivan et al., 2008: USA. Sullivan et al., 2010: USA. The review was conducted by authors based in the UK. **Source of funding:** Not reported.	(mean 7 years 6 months). Classified as middle childhood by the review authors. Bowen and Nowicki 2007: 7.5–10.5 years (mean not reported). Classified as middle childhood by the review authors. Camras et al., 1983: 3 years 7 months to 6 years, 4 months (mean not reported). Classified as early childhood by the review authors. Camras et al. 1988: 3 years 4 months to 7 years 3 months (mean 4 years 11 months). Classified as early childhood by the review authors. During and McMahon 1991: 2 years 8 months to 9 years 7 months (mean 5 years 8 months). Classified as early childhood by the review authors. Edwards et al. 2005: 5–12 years (mean 9 years 2 months). Classified as middle childhood by the review authors. Gapen 2010: 18–74 years (mean not reported). Classified as adolescence and adulthood by the review authors. Leist and Dadds 2009: 16–18 years (mean not reported). Classified as adolescence and adulthood by the review authors. Pajer et al. 2010: 16–18 years (mean not reported). Classified as adoles-	those measuring emotion knowledge was not significant (Q(1)=0.552; p=.457) but the difference between studies measuring emotion understanding and emotion recognition was (Q(1)=4.198; p=.040). The difference between studies measuring emotion knowledge and emotion recognition was also significant (Q(1)=10.873; p=.001). Meta-analysis – moderated by age group (negative effect size corresponds to poorer performance by maltreated participants) – A second moderator analysis examined the effect of age group. This found that studies with an 'early childhood' sample showed larger effect sizes than those with a 'middle childhood' sample, which in turn had larger effect sizes than those with an adolescence and adulthood sample. In fact, studies conducted in adolescence and adulthood showed a very small and non-significant effect. Early childhood: d=-0.933; 95% CI -1.160 to -0.706; Z=-8.065; p<.000; Q=11.320 (between studies); df (Q) 2; p=.003. Middle childhood: d=-0.776; 95% CI -1.315 to -0.236; Z=-2.818; p=.005. Adolescence and adulthood: d=0.042; 95% CI -0.479 to -0.563; Z =0.18; p=.875. Post-hoc comparison showed that these differences were significant for studies conducted in early childhood compared to those conducted in adolescence (Q(1)=11.320; p=.001) but not for early childhood compared to middle childhood (Q(1)=0.278; p=.598). The difference between studies conducted in middle childhood and adolescence was also significant (Q(1)=4.566; p=0.033).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	cence and adulthood by the review authors. Pears and Fisher 2005: 3–5 years (mean not reported). Classified as early childhood by the review authors. Perlman et al. 2008: 5–6 years (mean not reported). Classified as early childhood by the review authors. Pollak et al., 1997: 7.1–11.4 years (mean not reported). Classified as middle childhood by the review authors. Pollak et al. 2001: 6.3–12.2 years (mean 8.8 years). Classified as middle childhood by the review authors. Shackman and Pollak 2005: 7–12 years (mean 9.57 years). Classified as middle childhood by the review authors. Shipman and Zeman 1999: 6–12 years (mean not reported). Classified as middle childhood by the review authors. Shipman et al. 2005: 6–12 years (mean 9 years 3 months). Classified as middle childhood by the review authors. Smith and Walden 1999: 3 years 4 months to 6 years 0 months (mean not reported). Classified as early childhood by the review authors. Sullivan et al. 2008: 4–5 years (mean not reported). Classified as early childhood by the review authors. Sullivan et al. 2010: 4 years (range and mean	Effect sizes of individual studies included in the meta-analysis (calculated by review authors; Cohen's d; p values reported where provided, (negative effect size corresponds to poorer performance by maltreated participants) – 16 of the 19 studies (84.2%) showed effect sizes in the expected direction, that is maltreatment status or severity was associated with poorer emotion skills although only 12 of these were significant. Three of the 19 studies showed effect sizes in the reverse direction, that is maltreatment status or severity was associated with better emotion skills, although only one of these was significant. Barahal et al. 1981: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion recognition and emotion understanding, with a large effect size (controlled for IQ); d=-0.953; p=.010. Bowen and Nowicki, 2007: Participants in the maltreated group displayed significantly better performance than those in the non-maltreated group on measures of emotion recognition, with a very small effect size (controlled for IQ and language comprehension); d=0.078; p=.023. Camras et al., 1983: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion recognition, with a large effect size (not controlled); d=-1.058; p=.004. Camras et al., 1988: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion recognition, with a large effect size (not controlled); d=-1.018; p=.002. During and McMahon, 1991: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	not reported). Classified as early childhood by the review authors. Sex - The authors report that all studies had a mixed sample with the exception of Pajer et al. 2010; the sample of which was entirely female. Percentages are not reported by the review authors. • Ethnicity - Not reported for any of the included studies. • Religion/belief - Not reported for any of the included studies. • Disability - Not reported for any of the included studies. • Long term health condition - Not reported for any of the included studies. • Sexual orientation - Not reported for any of the included studies. • Socioeconomic position - Not reported for any of the included studies. • Type of abuse - The authors classified the abuse experienced by participants as physical abuse; sexual abuse; emotional abuse; physical neglect; domestic violence (witnessed); and verbal aggression (percentages are not reported). They also record whether the study compared a maltreated sample to a non-maltreated sample or used severity	group on measures of emotion recognition, with a large effect size (not controlled); d=-0.951; p=.002. Edwards et al., 2005: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion understanding, with a large effect size (not controlled); d=-2.902; p<.001. Gapen, 2010: No significant difference between maltreated and non-maltreated participants on measures of emotion recognition (not controlled, no statistical data reported, significance reported by review authors). Leist and Dadds, 2009: The correlation between prevalence or severity of maltreatment, and emotional recognition was non-significant (not controlled); d=0.448; p = .328. NB The positive effect size indicates that participants who had experienced greater levels of maltreatment performed better. Pajer et al., 2010: No significant difference between maltreated and non-maltreated participants on measures of emotion recognition, with a small to medium effect size (not controlled); d=-0.449; p=.084. Pears and Fisher, 2005: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion recognition and emotion understanding, with a large effect size (controlled for age and intelligence); d=-0.937; p<.001. Perlman et al., 2008: No significant difference between maltreated and non-maltreated participants on measures of emotion understanding, with a medium effect size (not controlled); d=-0.517; p<.133. Pollak et al., 1997: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion recognition, with a medium to large effect size (variables tested at baseline but no differences found in results); d=-0.707; p =	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	of maltreatment measured as a continuous variable. The details for the individual studies included in the meta-analysis were - Barahal et al. 1981: Physical abuse (maltreated vs. non-maltreated). Bowen and Nowicki 2007: Physical abuse, sexual abuse, emotional abuse, physical neglect (maltreated vs non-maltreated). Camras et al. 1983: Physical abuse (maltreated vs non-maltreated). Camras et al. 1988: Physical abuse (some physical neglect; maltreated vs non-maltreated). During and McMahon 1991: Physical abuse (maltreated vs non-maltreated). Edwards et al. 2005: Physical neglect (maltreated vs. non-maltreated). Gapen 2010: Physical abuse, sexual abuse, emotional abuse, physical neglect (severity of maltreatment measured as a continuous variable). Leist and Dadds 2009: Physical abuse, emotional abuse, physical neglect (severity of maltreatment measured as a continuous variable). Pajer et al. 2010: Physical abuse, sexual abuse, emotional abuse, physical neglect (maltreated vs. non-maltreated). Pears and Fisher, 2005: Physical abuse, sexual abuse, emotional	.023. Pollak et al., 2001: No significant difference between maltreated and non-maltreated participants on measures of emotion recognition, with a very small effect size (variables tested at baseline but no differences found in results); d=0.187; p=.569. NB The positive effect size indicates that participants who had experienced greater levels of maltreatment performed better. Shackman and Pollak, 2005: No significant difference between maltreated and non-maltreated participants on measures of emotion recognition, with a very small effect size (not controlled); d=-0.076; p=.764. Shipman and Zeman, 1999: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion understanding, with a large effect size (variables tested at baseline but no differences found in results); d=-1.405; p<.001. Shipman et al., 2005: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion understanding, with a medium to large effect size (variables tested at baseline but no differences found in results); d=-0.685; p=.021. Smith and Walden, 1999: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group relation to emotion recognition and emotion understanding, with a large effect size (controlled for 'receptive vocabulary'); d=0.863; p=.024. Sullivan et al., 2008: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of a composite measure of emotion recognition and emotion understanding, with a large effect size (controlled for IQ); d=-1.171; p=.003. Sullivan et	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	abuse, physical neglect (maltreated vs non-maltreated). Perlman et al. 2008: Physical abuse (maltreated vs. non-maltreated). Pollak et al. 1997: Physical abuse, emotional abuse, physical neglect (maltreated vs non-maltreated). Pollak et al. 2001: Physical abuse, physical neglect (maltreated vs. non-maltreated). Shackman and Pollak 2005: Physical abuse (maltreated vs. non-maltreated). Shipman and Zeman 1999: Physical abuse (maltreated vs non-maltreated). Shipman et al., 2005: Physical neglect (maltreated vs. non-maltreated). Smith and Walden 1999: Physical abuse, sexual abuse, physical neglect, witnessed domestic violence (maltreated vs non-maltreated). Sullivan et al. 2008: Physical neglect (some physical abuse, maltreated vs. non-maltreated). Sullivan et al. 2010: Physical neglect (maltreated vs. non-maltreated). • Looked after or adopted status Not reported for any of the included studies; however the authors report that studies which did not report pre-adoption histories were excluded suggesting that some of the studies included	al., 2010: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of a composite measure of emotion recognition and emotion understanding, with a large effect size (controlled for IQ); d=-1.008; p<.001. NB. Where studies measured both emotion recognition and emotion understanding the review authors calculated an average effect size. However, there were some studies which only reported scores on the composite measure of emotion knowledge (recognition and understanding). **Narrative findings** Meta-analysis - overall – (negative effect size corresponds to poorer performance by maltreated participants) – 16 of the 19 studies (84.2%) showed effect sizes in the expected direction, that is maltreatment status or severity was associated with poorer emotion skills although only 12 of these were significant. Three of the 19 studies showed effect sizes in the reverse direction, that is maltreatment status or severity was associated with better emotion skills, although only one of these was significant. The overall mean effect size across the 19 studies showed a medium effect size in the direction of maltreated children showing poorer emotion skills. Meta-analysis – moderated by outcome variable (negative effect size corresponds to poorer performance by maltreated participants) – A moderator analysis examining the effect of choice of outcome variable was conducted. The results suggested that the type of outcome measure did moderate the findings	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	in the meta-analysis may have included children who had been adopted or fostered. On the basis of the titles of included studies it appears that the sample assessed in Pears and Fisher 2005; included children in foster care. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported for any of the included studies. **Sample size:** • The studies included in the meta-analysis gave a combined total sample of 6,155 participants. The sample sizes of studies included in the meta-analysis were - Barahal et al. 1981: 17 maltreated vs. 16 non-maltreated participants. Bowen and Nowicki 2007: 1068 maltreated vs. 4166 non-maltreated participants. Camras et al. 1983: 17 maltreated vs. 17 non-maltreated participants. Camras et al. 1988: 20 maltreated vs. 20 non-maltreated participants. During and McMahon 1991: 23 maltreated vs. 23 non-maltreated participants. Edwards et al. 2005: 24 maltreated vs 24 non-maltreated participants. Gapen 2010: 162	(Q(2)=13.001, p=0.002), with studies measuring emotion understanding showing larger effect sizes than those measuring composite emotion knowledge, which in turn were larger than those measuring emotion recognition: The authors note that this may be because emotion understanding is a more advanced skill, and so may be 'particularly susceptible to the deleterious effects of maltreatment experiences' (p20). Post-hoc comparisons showed that the difference between studies measuring emotion understanding and those measuring emotion knowledge was not significant but the difference between studies measuring emotion understanding and emotion recognition was. The difference between studies measuring emotion knowledge and emotion recognition was also significant. Meta-analysis – moderated by age group (negative effect size corresponds to poorer performance by maltreated participants) – A second moderator analysis examined the effect of age group. This found that studies with an 'early childhood' sample showed larger effect sizes than those with a 'middle childhood' sample, which in turn had larger effect sizes than those with an adolescence and adulthood sample. In fact, studies conducted in adolescence and adulthood showed a very small and non-significant effect: Post-hoc comparison showed that these differences were significant for studies conducted in early childhood compared to those conducted in adolescence but not for early childhood compared to middle childhood. The difference between studies conducted in middle childhood and adolescence was also significant.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	participants (severity of maltreatment measured as a continuous variable). Leist and Dadds 2009: 23 participants (severity of maltreatment measured as a continuous variable). Pajer et al. 2010: 41 maltreated vs. 24 non-maltreated participants. Pears and Fisher 2005: 60 maltreated vs. 31 non-maltreated participants. Perlman et al. 2008: 17 maltreated vs. 18 non-maltreated participants. Pollak et al. 1997: 23 maltreated vs. 21 non-maltreated participants. Pollak et al. 2001: 28 maltreated vs. 14 non-maltreated participants. Shackman and Pollak 2005: 33 maltreated vs. 30 non-maltreated participants. Shipman and Zeman, 1999: 22 maltreated vs. 22 non-maltreated participants. Shipman et al. 2005: 24 maltreated vs. 24 non-maltreated participants. Smith and Walden 1999: 15 maltreated vs 15 non-maltreated participants. Sullivan et al., 2008: 12 maltreated vs 19 non-maltreated participants. Sullivan et al. 2010: 15 maltreated vs 27 non-maltreated participants. • Systematic reviews: number of studies - 19 studies were included in the meta-analysis.	Effect sizes of individual studies included in the meta-analysis (calculated by review authors, (negative effect size corresponds to poorer performance by maltreated participants) – 16 of the 19 studies (84.2%) showed effect sizes in the expected direction, that is maltreatment status or severity was associated with poorer emotion skills although only 12 of these were significant. Three of the 19 studies showed effect sizes in the reverse direction, that is maltreatment status or severity was associated with better emotion skills, although only one of these was significant. Barahal et al., 1981: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion recognition and emotion understanding, with a large effect size (controlled for IQ). Bowen and Nowicki, 2007: Participants in the maltreated group displayed significantly better performance than those in the non-maltreated group in relation emotion recognition, with a very small effect size (controlled for IQ and language comprehension); d=0.078; p=.023. Camras et al., 1983: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion recognition, with a large effect size (not controlled). Camras et al., 1988: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion recognition, with a large effect size (not controlled). During and McMahon, 1991: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion recognition, with a large effect size (not controlled). Edwards et	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Recognition indicators measured: Emotional understanding. The meta-analysis measures the effects of maltreatment on emotion skills only (due to methodological concerns and a paucity of studies measuring other social understanding related skills). The authors classify these as emotion recognition; emotion understanding; or emotion knowledge (a composite of emotion understanding or emotion knowledge). Where a study measured both emotion recognition and emotion understanding, an effect size was calculated for each of these and then '… averaged to produce an overall effect size for the sample' (p13). Barahal et al. 1981: Measured emotion recognition (labelling emotions) and emotion understanding (identifying causes of emotions). Bowen and Nowicki 2007: Measured emotion recognition (labelling emotions). Camras et al., 1983: Measured emotion recognition (choosing a photo for emotion label). Camras et al. 1988: Measured emotion recognition (choosing a photo for emotion label). During and McMahon 1991: Measured emotion recognition (labelling emo-	al., 2005: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion understanding, with a large effect size. Gapen, 2010: No significant difference between maltreated and non-maltreated participants on measures of emotion recognition (not controlled, no statistical data reported, significance reported by review authors). Leist and Dadds, 2009: The correlation between prevalence or severity of maltreatment, and emotional recognition was non-significant (not controlled). NB The effect size was positive indicating that participants who had experienced greater levels of maltreatment performed better. Pajer et al., 2010: No significant difference between maltreated and non-maltreated participants on measures of emotion recognition (not controlled). Pears and Fisher, 2005: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion recognition and emotion understanding, with a large effect size (controlled for age and intelligence). Perlman et al., 2008: No significant difference between maltreated and non-maltreated participants on measures of emotion understanding (not controlled). Pollak et al., 1997: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion recognition, with a medium to large effect size (variables tested at baseline but no differences found in results). Pollak et al., 2001: No significant difference between maltreated and non-maltreated participants on measures of emotion recognition (variables tested at baseline but no differences found in results). NB The effect size was	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	tions). Edwards et al. 2005: Measured emotion understanding (interview: '… identifying emotions and understanding causes, expression and consequences') (p10). Gapen 2010: Measured emotion recognition (labelling emotions). Leist and Dadds 2009: Measured emotion recognition (labelling emotions). Pajer et al. 2010: Measured emotion recognition (labelling emotions). Pears and Fisher 2005: Measured emotion recognition and emotion understanding ('… labelling emotions, selecting photo for emotion label and emotion for story character') (p10) Perlman et al. 2008: Measured emotion understanding ('Rating plausibility of pairing of emotion and cause') (p10). Pollak et al. 1997: Measured emotion recognition ('Describing target emotion face from display') (p 10). Pollak et al. 2001: Measured emotion recognition ('Describing target emotion face from display') (p10). Shackman and Pollak, 2005: Measured emotion recognition (labelling emotions). Shipman and Zeman 1999: Measured emotion understanding (interview: '… identifying emotions and understanding causes, expression and consequences') (p10). Shipman et	positive indicating that participants who had experienced greater levels of maltreatment performed better. Shackman and Pollak, 2005: No significant difference between maltreated and non-maltreated participants on measures of emotion recognition (not controlled). Shipman and Zeman, 1999: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion understanding, with a large effect size (variables tested at baseline but no differences found in results). Shipman et al., 2005: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion understanding, with a medium to large effect size (variables tested at baseline but no differences found in results). Smith and Walden, 1999: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of emotion recognition and emotion understanding, with a large effect size (controlled for 'receptive vocabulary'). Sullivan et al., 2008: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of a composite measure of emotion recognition and emotion understanding, with a large effect size (controlled for IQ). Sullivan et al., 2010: Participants in the maltreated group displayed significantly poorer performance than those in the non-maltreated group on measures of a composite measure of emotion recognition and emotion understanding, with a large effect size (controlled for IQ).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	al. 2005: Measured emotion understanding (interview: '… identifying emotions and understanding causes, expression and consequences') (p10). Smith and Walden, 1999: Measured emotion recognition and emotion understanding ('Selecting similar emotion face to target and for story character') (p10). Sullivan et al. 2008: Measured and reported on a composite measure of emotion knowledge (recognition and understanding – '… labelling emotions, selecting photo for emotion label and emotion for story character …') (p10). Sullivan et al. 2010: Measured and reported on a composite measure of emotion knowledge (recognition and understanding – '… labelling emotions, selecting photo for emotion label and emotion for story character …' (p10).		

13. Miller Adam B et al. (2013) The relation between child maltreatment and adolescent suicidal behavior: A systematic review and critical examination of the literature. Clinical Child and Family Psychology Review 16: 146–72

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To review literature to date examining the relationship between CM (child mal-	**Participants:** Children and young people. **Sample characteristics:**	**Statistical data -** No quantitative data reported in the study. See narrative findings. **Narrative findings -** A. Sexual abuse 1. Cross-sectional studies of community samples (n=28; mean sample size = 6,177	**Overall assessment of internal validity:** + **Overall assessment of external validity;** ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
treatment) and adolescent suicidal ideation and attempts. **Methodology:** Systematic review. **Country:** Range of countries. USA (13 studies), New Zealand (4 studies), Switzerland (2 studies), Canada (2 studies), Brazil (1 study), Netherlands (1 study), Italy (1 study), Australia (1 study), France (1 study), country not reported (26 studies). **Source of funding:** Government – A federal grant (R01AA016854) and in part by NIAAA grant R01AA016854-04.	• Age - Studies were included if the sample was composed 'predominantly' of adolescents aged 12–17 years. For some studies, the upper age limit was higher than this. • Sex - Studies range from 35 to 88% females. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Physical abuse, neglect, majority included studies on sexual abuse (n=47). • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - Control groups in some studies but not in all studies. Intervention numbers - NA (x-sectional studies and longitudinal studies).	participants per study; ranged from 100 to over 80,000), using a variety of methods to assess sexual abuse history and suicidal ideation and suicide attempts; 27 of the 28 studies 'demonstrated clear evidence of a general association between a history of sexual abuse and increased suicidal ideation and/or suicide attempts' (p3). One study with small sample size failed to find this link (Arata 2007). 2. Cross-sectional studies of clinical/high-risk populations (psychiatric inpatients, delinquent youth) (n=16; mean sample size=468 participants per study, range from 48 to 2,019); 14 of 16 studies that focused on clinical/high-risk populations also found an association between childhood sexual abuse and adolescent suicidal ideation and suicide attempts. Two studies with small sample size and a control group did not find this association. 3. Longitudinal studies of community samples (n=8; mean sample size=594 participants per study; range from 133 to 1,631; the time period for the follow-up between 6 months to 28 years). All showed evidence to suggest that childhood sexual abuse predicts future suicidal ideation and/or suicide attempts in adolescence. 4. Effect of covariates The association between sexual abuse and suicidal behaviour/ideation remains significant when controlling for demographic variables of age and grade level (11 studies); sex (8 studies), IQ (1 study) and race/ethnicity (4 studies); youth mental health problems (7 studies); general psychiatric symptoms during childhood and early adolescence (1 study); family structure (2 studies); parental separation (1 study); mothers' level of education (1 study); family socioeconomic status (4 studies); parental violence or imprisonment (1 study); parenting style or family functioning (3 studies); parents' psychiatric symptoms and substance abuse (3	**Overall validity rating:** + Of moderate quality in terms of a narrative review of studies. Unclear why statistical data not reported, or statistical analyses not conducted.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Systematic reviews: number of studies - 29 studies on child maltreatment and suicidal ideation (18 cross-sectional studies on community samples; 9 cross-sectional studies on clinical/high risk samples; 1 longitudinal study on community samples; 1 longitudinal study on clinical/high risk samples); 49 studies on child maltreatment and suicidal attempt (25 cross-sectional studies on community samples; 15 cross-sectional studies on clinical/high risk samples; 8 longitudinal study on community samples; 1 longitudinal study on clinical/high risk samples) involving 182,018 adolescents. Included studies published between 1989 and 2012. **Recognition indicators measured:** Suicidal thoughts/behaviour and suicidal ideation and attempts.	studies) and parental suicide (1 study). The association is not clear when controlling for negative life events. There is some evidence that accumulative negative life events may affect the relationship between sexual abuse and suicidal ideation/suicide attempts. B. Physical Abuse 1. Cross-sectional studies conducted with community samples (n=18; mean sample size=3694 participants per study, range 1214 to 16,644). Sixteen of 18 studies showed a positive relationship between childhood physical abuse and suicidal ideation and/or attempts. 2 studies with small sample size did not show this association. 2. Cross-sectional studies of clinical and high-risk samples (psychiatric inpatients, delinquent youth) (n=10, mean sample sizes=499 participants per study, range 114 to 2,019). Nine out of 10 studies revealed a relationship between childhood physical abuse and suicidal ideation and/or attempts. One study with small sample size failed to find this link. 3. Longitudinal studies conducted with community samples (n=6, mean sample size=745 adolescents, range from 200 to 1,631; followed over periods of 6 months–28 years). Five out of 6 found an association between physical abuse and adolescent suicidal ideation and/or attempts. One study did not find this link. 4. Effects of covariates This positive association remains significant when controlling for age (9 studies), sex (9 studies), race/ethnicity (3 studies), family socioeconomic status (1 study), or caregiver education level (1 study). In 7/8 studies reviewed (4 cross-sectional studies and 3 longitudinal), there was no differences in these associations when controlling for psychological distress in	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		childhood and early adolescence (2 studies), depression severity (3 studies), disruptive and risky behavior (1 study) or prior suicide attempts (1 study). One study did find a change in this relationship. Similar positive association was found when controlling for a history of suicide within the family (1 study), family alcohol and drug problems (1 study), parent attachment (1 study) and parent psychiatric symptoms (1 study). C. Emotional Abuse and Neglect 1. Cross-sectional studies conducted with community samples (n=6), and clinically based sample (n=1) (mean sample size=845 participants per study; range from 114 to 2,247 participants). All found significant relationships between neglect and/or emotional abuse, and adolescent suicidal ideation or behaviour. 2. One 17-year longitudinal study (n=639 youth; 39 of which had substantiated cases of childhood neglect) found that childhood neglect did not predict future suicidal behaviour. 3. Effects of covariates: There were mixed results when 4 out of the 7 cross-sectional studies controlled for basic covariates. The association remained when controlled for sex in 2 of 4 studies reviewed. Emotional abuse, but not neglect, was found to be independently associated with suicide ideation after controlling for youth sex and race, youth mental health problems and family variables (1 study). In one longitudinal study (Brown et al. 1999) neglect did not predict suicide attempts after 21 potential risk factors were controlled. D. Co-occurrence of Sexual and Physical Abuse 1. 18 studies (4 longitudinal) showed mixed findings when both sexual and physical abuse were examined. Sex-	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		ual abuse and physical abuse were independently associated with suicidal ideation (3 studies) and/or suicide attempts (8 studies), after controlling for age (4 studies); sex (3 studies); ethnicity (1 study); socioeconomic status (1 study); psychiatric symptoms and diagnoses (2 studies); prior suicide attempts (1 study), parental psychiatric symptoms (1 study), and family alcohol and drug use problems (1 study). 2. When sexual abuse and physical abuse were examined simultaneously, only sexual abuse was associated with various measures of suicidal ideation and behavior (4 studies), after controlling for socioeconomic status (2 studies), youth dissociative symptoms (1 study), youth negative life events (1 study); parental violence, parental mental health symptoms, parental imprisonment (1 study), mother's education, parenting etc. (1 study). 3. There was an additive effect of sexual and physical abuse on suicide attempts (3 studies). Youth victims of both forms of abuse were more likely to report suicide attempts (3 studies) than either alone, as well those with no abuse (1 study), both in any suicide attempt (3 studies) as well as multiple attempts (1 study), the latter only found for females. One study showed an additive effect of both forms of abuse on suicidal ideation, and 1 study did not. E. All forms of child maltreatment (Sexual Abuse, Physical Abuse, Emotional Abuse and Neglect) The review undertook a multivariate analysis of the relative contribution of each form of child maltreatment (sexual abuse, physical abuse, emotional abuse and neglect) to adolescent suicidal ideation and behaviour. 1. Thirteen studies examined this relationship. All forms of abuse were independently associated	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		with suicide attempts (5 studies) and/or suicidal ideation (2 studies). When controlled for contextual risk factors (sex, ethnicity, IQ, temperament, serious mental illness, anger, dissatisfaction, external locus of control, sociopathy, low religious participation, teenage pregnancy, single parenthood, welfare support, low family income, large family size, maternal factors, paternal factors, only sexual and physical abuse, not neglect, remained significant (1 study). 2. Only sexual and emotional abuse remain significant when suicide attempts were the outcome variable (3 studies). Emotional abuse had the strongest association with suicide attempts, with no significant independent prediction by either physical abuse or neglect (1 study). After controlling for adverse childhood experience, only sexual abuse and low parental care, not physical abuse, were associated with suicide attempts (1 study). Only emotional abuse was associated with suicide attempts in 9th graders while sexual abuse was associated with suicide attempts in a combined sample of 9th and 11th graders. Physical abuse was not associated with suicide attempts in either grade (1 study). 3. Only physical and emotional abuse (not sexual abuse or neglect) were independently associated with suicidal ideation (1 study). Another study found that neglect remained significantly associated with suicide-related behaviours. 4. One study results suggest that risk of a suicide attempt increases with the addition of each form of abuse, providing evidence for an additive effect. F. Potential Moderators and Mediators 1. Sexual abuse was independently associated with suicide attempts for males only (7 studies), suggesting that childhood sexual abuse may be more strongly linked	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		to suicide attempts in males than in females. 2. There was a stronger association between childhood physical abuse and adolescent suicide attempts among males than females (2 studies). 3. There was a more consistent association between childhood physical abuse and adolescent suicidal ideation among females relative to males (2 studies). 4. There was no sex differences in the relationship between childhood physical abuse and adolescent suicidal ideation (2 studies). G. Characteristics of the Abuse Experience 1. The relative risk of a suicide attempt increased as a function of the severity of sexual abuse (1 study). Frequency of abuse episodes and sexual abuse experiences that involved contact (i.e., touching, intercourse) with the perpetrator were more likely to report a history of suicide attempts compared to adolescents who reported non-contact sexual abuse (i.e., verbal sexual harassment) or no sexual abuse history. 2. Other factors specific to the sexual abuse experience that were related to an increased risk of making a suicide attempt: a later age of onset of sexual abuse; when the perpetrator was an acquaintance (rather than an authority figure or caregiver); when a parent denied the abuse occurrence; when a parent expressed anger for the abuse incident toward the child rather than the perpetrator; and a history of a single episode of sexual abuse (1 study). 3. Only physical abuse from fathers predicted suicidal ideation in adolescent boys, whereas physical abuse from either parent predicted suicidal ideation in adolescent girls (1 study).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		H. Inter- and Intra-personal Factors 1. Among youth who reported more severe sexual abuse, those with higher levels of satisfaction with their social supports (from peers, family, and others) reported lower levels of suicidal ideation than those with more severe sexual abuse and lower satisfaction with their social supports (1 study). 2. Satisfaction with social support did not moderate the association between physical abuse severity and suicidal ideation (1 study). 3. Youth's problem-solving confidence moderated the association between physical abuse and suicidal ideation (1 study).	

14. Mironova P et al. (2011) Childhood physical abuse and suicide-related behavior: A systematic review. Vulnerable Children and Youth Studies 6: 1–7

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To investigate how shared environment with perpetrator(s) identified as a family member or parent/parental figure or an adult at home contribute to the association between childhood physical abuse and suicide-related behavior. **Methodology:** Systematic review. Exclusion criteria: Case reports, qualitative studies, reviews, and editorials, studies not reporting on	**Participants:** Children and young people – Children were school- or population-based, aged 18 years or younger or in grade 12 or less. (Perpetrator(s) had to be identified as a family member or a parent/parental figure or an adult at home.) **Sample characteristics:** • Age - up to 18 years. • Sex - average up to 44–61% males. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported.	**Statistical data** Association between childhood physical abuse (CPA) by a known perpetrator in the family and suicide-related behaviours (SRB) (3 cross-sectional studies): 1. Flisher et al. 1996: Unadjusted data - Significant association between CPA and SRB in both boys and girls. Adjusted data - Boys: Odds Ratio (OR)=2.4 (95% CI 1.3 to 4.5). Girls – no data reported 2. Lau et al. 2003: Unadjusted data: Significant association between CPA and SRB in both boys and girls. For corporal punishment, OR=3.3, p=0.054; for 'Beaten for no reason', OR=2.4 p=0.074; for 'Beaten to injury', OR=8.5, p<0.001. No adjusted effect reported 3. Logan et al. 2009: Significant association between CPA and SRB in both boys and girls. PR (prevalence ratio)=3.7, p<0.002 for girls; PR=2.4, p<0.05 for boys. Adjusted for child sexual abuse (CSA), sex, age, and	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** + Good reporting of methodology, and justification for lack of critical appraisal of included studies.

the CSA–SRB association as specified, studies in clinical or child welfare settings were also excluded. As described in Rhodes, A., Boyle, M., Tonmyr, L., Wekerle, C., Goodman, D., Leslie, B., Mironova, P., Bethell, J., Manion, I. (2011) Sex differences in childhood sexual abuse and suicide-related behaviors. Suicide & Life-Threatening Behaviour. **Country:** Range of countries. Studies conducted in South Africa (1 study), Hong Kong (1 study), USA (1 study), New Zealand (1 study), Canada (1 study). **Source of funding:** Government – Canadian Institutes of Health.	• Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Physical abuse. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** Systematic reviews: number of studies - 5 studies (3 X-sectional and 2 longitudinal studies, 12,262 children involved, sample sizes ranging from 489 to 7340). All questions were self-reported. **Recognition indicators measured:** • Internalising and externalising behaviour. • Suicide-related behaviours.	race and witnessing family violence: PR=2.5 (95% CI 1/.9 to 3.3) (1 study). Association between childhood physical abuse (CPA) by a known perpetuator in the family and suicide-related behaviours (SRB) Two longitudinal studies: 1. Fergusson and Lynskey, 1997: Significant association between CPA and SRB in young adults, no odds ratios reported. Adjusted for CSA, sex, family life events etc., rate of suicidal attempts by extent of physical punishment: none: 3.3%; seldom: 5%; regular: 7.4%; severe/harsh: 10.7%; p<0.05., no odds ratios reported. 2. Brezo et al. 2008: Significant association between CPA and SRB in young adults. Unadjusted data: CPA (no CSA) OR=1.8 (95% CI 1.1-3.0); CSA (no CPA) OR 2.0 (95% CI 1.1 to 3.9), CPA and CSA 4.6 (2.7 to 8.1) Adjusted for sex, disruptive behaviours etc., for CPA (no CSA), OR=1.9 (95% CI 1.0 to 3.6), p<0.05; CSA (no CPA): no data reported; CPA and CSA, OR 4.7 (95% CI 2.5 to 8.9), p<0.001. **Narrative findings** Unadjusted data from all five studies (Brezo et al. 2008; Fergusson and Lynskey 1997; Flisher et al. 1996; Lau et al. 2003; Logan et al. 2009) found statistically significant associations between physical abuse perpetrated by a family member and suicide-related behaviours. Three studies reported unadjusted odds ratios/prevalence ratios, which ranged from 1.8 (95% Ci 1.1 to 3.9) to 3.7 (95% CI not reported). Three studies reported adjusted odds ratios/prevalence ratios, after controlling for factors such as age, race and family violence, which ranged from 1.9 (95% CI 1.0 to 3.6) to 2.5 (95% CI 1.9 to 3.3). One of the included studies (Fergusson and Lynskey 1997) found that rates of suicide attempt in-	

		creased depending on the severity of physical punishment (adjusted significance level, p<0.05; no odds ratios reported). One study (Brezo et al. 2008) also examined the relationship between a combination of physical abuse and sexual abuse and suicide-related behaviours, estimating an adjusted odds ratio of 4.7 (95% CI 2.5-8.9).	

15. Naughton AM et al. (2013) Emotional, behavioral, and developmental features indicative of neglect or emotional abuse in preschool children: a systematic review. JAMA Pediatrics 167: 769–75

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To review evidence to 'define the emotional, behavioral and developmental features of neglect or emotional abuse in preschoolers' (p769). **Methodology:** Systematic review. **Country:** Range of countries. 40 included studies carried out in US, 2 in Canada. Reviewing team based in the UK. **Source of funding:** Other – UK National Society for the Prevention of Cruelty to Children.	**Participants:** • Children and young people - 35 of the 42 included studies related to children (aged 0–6 years) experiencing neglect, emotional abuse or emotional neglect. • Caregivers and families - 14 of the 42 studies related to parents and caregivers of children aged under 6. **Sample characteristics:** • Age - Of the 35 studies relating to children and young people: 8 related to children aged 0 to 20 months 5 related to children aged 20 to 30 months 5 related to children aged 3 to 4 years 6 related to children aged 4 to 5 years 5 related to children aged 5 to 6 years 6 were prospective cohort studies that followed up children at different ages. • Sex – Not reported. • Ethnicity - Not reported.	**Statistical data** NB - Data are taken from both narrative summary and data tables. As for the systematic review, we have grouped the findings by age group of children, and by the types of features explored in the study. This study reports p values only – no effect sizes or odds ratios are reported. 1. Child features The study states that 'emotional, behavioral and developmental features in the child associated with neglect or emotional abuse were described in 22 case control studies, 1 cross-section study and 12 cohort studies' (p770). <u>1.1 Children aged 0–20 months</u> 1.1.1 Attachment status. Three studies measured attachment of 1 year-old children using the Strange Situation task (Cicchetti et al. 2006; Crittenden 1985; Lamb et al. 1985). (Note that the data tables suggested that this was also examined in an additional study, Egeland 1981, but this is not reported in the systematic review). The studies found that a high proportion of maltreated infants were classified as disorganised on strange situation (89.8%, Cichetti et al. 2006), although comparable figure for non-maltreated	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall validity rating:** - Key limitations: Critical appraisal conducted but unclear how this was used within the analysis, unclear how study results were combined to arrive at the lists of indicators presented, poor reporting of statistical data from the original studies.

	• Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Studies were of children experiencing neglect or emotional abuse. Comparison groups were either children experiencing other forms of abuse or non-maltreated controls. Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** Systematic reviews: number of studies - 42 studies included. Total sample sizes for each study were not always clearly reported. These figures are based on information provided in the data tables: Allen 1982 USA 79 Cheatham 2010 USA 151 Christopoulos 1988 USA 30 Cicchetti 2006 USA 189 Crittenden 1984 60 Crittenden 1985 USA 121 Crittenden 1985b USA 77 Crittenden 1988 USA 105 Crittenden 1992 USA 182 Culp 1991 USA 74 DiLalla 1990 USA 120 Dubowitz 2002 USA 136 Dubowitz 2004 USA 173 Egeland & Sroufe 1981	infants not reported in the SR - neglected infants showed insecure avoidant status (Crittenden 1985), although again it is not clear how this compares to non-maltreated infants - 63% of neglected infants showed insecure avoidant attachment compared to controls (p<0.005) (Lamb et al. 1985). 1.1.2 Behaviour patterns with caregivers This was examined for this age group in 2 studies (Crittenden 1985; Crittenden and DiLalla 1988). The studies found that neglected infants had predominantly passive behaviour pattern of interaction with their mothers (12 out of 20) (p<0.001), although it is not clear how this compared to controls (Crittenden 1985). Neglected children were more passive initially but as they became older (12 months onwards up to 2 and a quarter) their negative and resistance behaviour increased. Neglected children with unresponsive mothers learned to display their anger rather than inhibit it in comparison with abused children (p<0.001) (Crittenden and DiLalla 1988). 1.1.3 Cognitive skills One study looked at the impact of abuse on cognitive skills in this age group (Mackner et al. 1997). This study found that the cognitive performance of the group with a combination of neglect and failure to thrive (FTT) was significantly below that of the children in the neglect only group (p<0.01), FTT only (p<0.01) and comparison group (p<0.01). Another study (Sylvestre and Merette 2010) examined language delay and found that language development delay in neglected children was more likely in children of depressed mothers (p=0.08). 1.1.4 Social interactions. One study (Valentino et al. 2006) examined the impact of maltreatment on infant social interactions and found no difference from nonmaltreated controls on complexity of play style or cognitive play abilities.	

	64 Egeland & Sroufe 1981 147 Egeland 1983 USA 147 Eigsti 2004 USA 33 English 2005 USA 212 Erickson 1989 USA 125 Fagan 1993 USA 27 Frodi 1984 USA 60 Hoffman-Plotkin 1984 USA 42 Koenig 2000 USA 115 Koenig 2004 USA 82 Lamb 1985 USA 48 Macfie 1999 USA 107 Macfie 2001 USA 258 Mackner 1997 USA 177 Maughan 2002 USA 88 Mustillo 2011 USA 573 Pianta 1989 USA 147 Pollak 2000 USA 48 Pollitt 1975 USA 29 Rohrbeck 1986 USA 38 Scarborough 2009 USA 997 Sullivan 2008 USA 31 Sylvestre 2010 Canada 68 Toth 1997 USA 107 Toth 2000 USA 69 Valentino 2006 USA 99 Venet 2007 Canada 74 Waldinger 2001 USA 31. **Recognition indicators measured:** • Attachment - Included studies use Strange Situation (Ainsworth et al. 1978). • Cognitive skills - Included studies used scales such as Stanford Binet Intelligence Scale. • Emotional understanding. • Language development - Included studies use tools include preschool language scale manual (Zimmerman et al. 1979), index of productive syntax (Scarborough 1990), Peabody picture	<u>1.2 Children aged 20 to 30 months</u> 1.2.1 Attachment status. Two studies examined attachment in this age group (Crittenden 1985, 1992) but no statistical data are reported. 1.2.2 Behaviour patterns with caregivers. One study (Crittenden 1992) found that neglected children spent least time with adults ($p<0.02$) and most time alone ($p<0.005$) compared to those experiencing other forms of abuse, those experiencing 'marginal' maltreatment and a control group. They also showed the most passive behaviour with their mothers ($p<0.001$). 1.2.3 Cognitive skills. One study (Cheatham et al. 2010) found that neglected children experienced deficits in performance on memory testing in comparison to abused and matched controls ($p<0.001$). 1.2.4 Social interactions. One study(DiLalla and Crittenden 1990) found that neglected children had significantly less positive social interaction compared with controls ($p<0.001$). <u>1.3 Children aged 3 to 4 years</u> 1.3.1 Attachment status. No studies examined this measure in this age group. 1.3.2 Behaviour patterns with caregivers. One study (Koenig et al. 2000) found that neglected children demonstrated significantly more negative affect (anger) than either physically abused or non-maltreated children. No statistical data reported. 1.3.3 Cognitive skills. Two studies examined language abilities (Allen and Oliver 1982; Culp et al. 1991). They found that neglected children had reduced comprehension and expressive language abilities compared to abused children and control groups ($p>0.001$) (Allen and Oliver 1982) and that, compared to children experiencing physical abuse, those experiencing neglect had the lowest scores on auditory and verbal scores on the preschool language scale ($p<0.01$) and lowest scores on profile language subscale ($p<0.01$). Neglect was most strongly associated with expressive	

	vocabulary test (revised Dunn 1981). • Play - Included studies used scales such as cheating game (Kochanska 1996), child structural play scale (Nicolich 1977). • Social interactions/peer relationships - Included studies used scales such as Rothenberg social sensitivity test (Rothenberg 1970), emotion recognition task (adapted by Ribordy 1988), emotional discrimination task (Borod 1990). • Internalising and externalising behaviour. • Self-esteem.	and receptive language delay. 1.3.4 Social interactions. One study (Frodi and Smetana 1984) found that, in children younger than 4 years, when IQ entered as a co-variate, there no difference between groups on their ability to discriminate emotion in others (no statistical data given). 1.4 Children aged 4 to 5 years 1.4.1 Attachment status. No studies found measuring this indicator for this age group. 1.4.2 Behaviour patterns with caregivers. No studies found measuring this indicator for this age group. 1.4.3 Cognitive skills. One study (Eigsti and Cicchetti 2004) found that maltreated children showed a 16 month delay in syntactic development for language compared with 13 months for controls. Scores on Peabody picture vocabulary test were lower in maltreated groups compared with controls ($p<0.04$). A second study (Hoffman-Plotkin and Twentyman 1984) found that neglected children had lower scores on cognitive functioning compared with non-maltreated controls ($p<0.01$) 1.4.4 Social interactions. Four studies examined social interactions. One study (Hoffman-Plotkin and Twentyman 1984) found that neglected children neglected children engaged in the least number of interactions with other children, especially prosocial behaviour ($p<0.05$), and also showed more disruptive behaviour ($p<0.01$). One study (Macfie et al. 1999) found that neglected children portrayed the children (story stem) as responding less often to relieve distress in other children (no statistical data provided) and a second study (Pollak et al. 2000) found that neglected children showed less accurate recognition of anger ($p<0.05$) and significantly less recognition of disgust than controls ($p<0.01$). One study comparing neglected children with those who had suffered other forms of abuse and a control group (Rohrbeck and Twentyman 1986) found that	

		mean parent and teacher ratings of neglected children were in the direction of greater dysfunction than either the abuse or control children (p<0.05). 1.5 Children aged 5 to 6 years 1.5.1 Attachment status. One study (Venet et al. 2007) found that neglected children displayed more overall disorganised markers (p<0.01) than a control group. Neglected children showed a significant difference in attachment representations even when socioeconomic status and maternal stress were controlled (p<0.05). Neglected group had a significantly higher proportion of avoidant attachment classification compared to controls (p<0.01). 1.5.2 Behaviour patterns with caregivers. One study (Venet et al. 2007) found that neglected children depicted their mothers as being absent or less available compared with controls (p<0.05). 1.5.3 Language and cognitive skills. No studies found measuring this indicator for this age group. 1.5.4 Social interactions. One study (Waldinger et al. 2001) found that neglected children were more likely to represent another child as sad, hurt or anxious (p<0.01) and more likely to see self as shamed or anxious (p=0.06) compared to control children. This study also found that neglected children represent themselves as angry and opposing others more frequently than non-maltreating (p<0.05). A second study (Toth et al. 1997) found that neglected children had the lowest positive self-representation compared with sexual abuse, physical abuse and controls (p<0.01). Neglected children had more negative maternal representations compared with controls, though not as marked as physical abuse children (p<0.001). Another study (Koenig et al. 2004) found that neglected children engaged in significantly more cheating behaviour (p<0.01) and less rule compatible behaviour (p<0.05). Another study found that peer relationships	

		were affected by increased rates of dissociation (Mafie et al. 2001). 1.6 Transition through development Eleven studies followed up neglected or emotionally abused children through their childhoods, covering a total of 1626 cases. 1.6.1 Attachment status. This was measured in two studies. One study (Egeland and Sroufe 1981) found that 50% of neglected children were rated as attachment type C (insecure, ambivalent or resistant) at 1 month compared to only 9% of control (p=0.008). At 18 months the neglect group shifted from a C type to A (insecure avoidant) (37% as opposed to 14% at 12 months) types but some also classified as secure (47% as opposed to 36% at 12 months). In a second study (Egeland and Sroufe 1981b) found a marked increase in maladaptive patterns of functioning. 1.6.2 Behaviour patterns with caregivers. Crittenden (1985) found that neglected infants were passive initially, but from 12 months onwards shoed aggressive and resistance behaviour towards caregivers (no statistical data reported). 1.6.3 Language and cognitive skills. One study with 212 participants (English et al. 2005) found that impairment in expressive language associated with dirty, unsafe residence (p<0.001) and failure to provide shelter. Receptive language impairment was associated with untreated emotional or behavioural problems (p<0.0005). A second study (Erickson et al. 1989) of 125 children experiencing neglect, emotional neglect, physical abuse, sexual abuse and a non-abused control found that neglected children had lower scores on Wechsler Preschool and Primary Scale of Intelligence (WIPPSI) for comprehension, vocabulary and animal house subtests as well as total tests compared with controls, emotional neglect and	

		sexual abuse groups. Teachers rated neglected children as having more difficulty comprehending school work ($p<0.01$), lacking creative initiative ($p<0.01$), poorer at following direction ($p<0.05$), low in reading and expressing themselves ($p<0.05$). Emotional neglected children had lower scores on WIPPSI though not as significant as neglect children. A third study found that neglected children showed more anger in problem solving task ($p<0.01$). Another study (Egeland and Sroufe 1981) found that in a study comparing children experiencing neglect, emotional neglect, emotional abuse and a non-abused control, the emotionally abused group had the greatest decrease in their cognitive functioning score from 9 to 12 months, whereas at 24 months children experiencing neglect showed the greatest anger when performing problem-solving tasks. A study following up children from 18 months to 3 years (Scarborough et al. 2009) found that developmental delay in preschool children was predicted by neglectful care (failure to provide for basic needs rather than failure to supervise) ($p<0.05$). 1.6.4 Social interactions. Four studies examined the impact of neglect on behaviour and social interactions. One study (Dubowitz et al. 2002) found that at age three, psychological neglect was significantly associated with children's internalising ($p<0.01$) and externalising behaviour problems ($p<0.001$). Neglect measured at age three did not predict changes in children's development and behaviour between ages 3–5. Cumulative neglect index was associated with internalising problems (depression/passivity ($p<0.001$). A second study (Dubowitz et al. 2004) found that ore difficulties with peer relations with psychological neglect, at age 6 on teacher report ($p<0.01$). A third study found that neglected children had early deficits in emotional knowledge across all three components of labelling ($p<0.01$), visual recognition and matching	

		context (Sullivan et al. 2008). A fourth longitudinal study found that at follow-up neglected children evidenced more negative self-representation than did non-maltreated children ($p<0.001$). 2. Carer-child interactions Fourteen studies addressed characteristics of carer-child interactions in cases of emotional abuse and neglect. 2.1 Children aged 0 to 12 months Two studies examined carer-child interactions in this age group. One study (Cicchetti et al. 2006) found that maltreating mothers reported infants as less reinforcing, accepting and adaptable and more demanding than those in a control group ($p<0.01$). Mothers in the maltreatment group were also rated as substantially lower in maternal sensitivity compared to non-maltreating mothers ($p<0.001$). A second study (Christopoulos et al. 1988) examined language patterns in neglecting and non-neglecting mothers. They found that mothers who neglected their infants used fewer commands ($p<0.01$). 2.2 Children aged 1 to 3 years Eight studies examined carer-child interactions for this age group. One study (Crittenden and DiLalla 1988) found that neglectful mothers were unresponsive, meaning that children learned to display their anger rather than inhibit it in comparison with abused children ($p<0.001$). A second study (Crittenden and Bonvillian 1984) found that neglecting mothers seemed to be withdrawn and uninvolved with their infants, they expressed little or no affection to their children and initiated few activities with them. Involvement in play was sporadic and minimal ($p<0.001$). A third study (DiLalla and Crittenden 1990) found that, compared to abused/neglected children and a control group, ne-	

		glectful parents showed the least positive social interaction with their children ($p<0.001$). Fagan and Dore (1993) found that, compared to a control group, neglectful mothers were less developmentally appropriate towards their children (less attuned) in free play ($p<0.05$). Neglecting mothers significantly less responsive to children than adequately rearing mothers ($p<0.01$) even when controlling for education level. Two studies (Koenig et al. 2000 and Mustillo et al. 2011) explored the link between maternal mood and neglect, and found that low maternal affect was linked to child depression. Another study (Pianta et al. 1989) explored interactions during a problem solving task. Psychologically unavailable mothers ignored their child's cues for assistance, offered no encouragement even if the child was failing and looked comfortable even when the child was highly frustrated ($p<0.05$). A final study (English et al. 2005) found that parental verbal aggression and verbally aggressive discipline was associated with child anxiety, depression ($p<0.001$) and attention problems ($p<0.001$). 2.3 Children aged 3 to 6 years Four studies explored carer-child interactions in this age group. One study with 33 participants (Eigsti et al. 2004) found that mothers in maltreating groups produced a small number of utterances ($p=0.025$), spoke less frequently ($p=0.04$) and produced fewer complex multi-clause utterances ($p=0.04$). Maternal verbal IQ scores correlated significantly with child Peabody picture vocabulary test scores ($p=0.02$) and child Index of productive syntax scores ($p=0.04$). A second study (Macfie et al. 1999) found that, using Story Stem vignettes, and found that neglected children portrayed their parent's s responding less often to relieve distress. A third study (Toth et al. 1997) found that neglected children	

		had more negative maternal representations compared with control children, though not as marked as those experiencing physical abuse ($p<0.001$). Finally, Pollitt et al. (1975) found that mothers of neglected children were less likely to relate to their children ($p<0.01$), were less affectionate in comparison with the control mothers, who used more positive verbal instruction, praise and positive contact with their children ($p<0.01$). Controls had twice as many recorded instances of positive affect than index cases. **Narrative findings** This systematic review concludes that the following emotional, behavioural and developmental features are indicative of neglect or emotional abuse: Child features 0 to 20 months - insecure-avoidant attachment (Crittenden 1985; Lamb et al. 1985); insecure-disorganized attachment (Cicchetti et al. 2006), cognitive skills and developmental delay (Mackner et al. 1997; Sylvestre & Merette 2010; Valentino et al. 2006), passive withdrawn behaviour (Crittenden and DiLalla 1988); 20–30 months - negativity in play (DiLalla and Crittenden 1990), reduced social interactions (Crittenden 1992) and deficits in memory performance (Cheatham et al. 2010). Three to 4 years - negativity in play (Koenig et al. 2000), delays in complex language (Allen and Oliver 1982; Culp et al. 1991); difficulties with emotion discrimination (Frodi and Smetana 1984). Four to 5 years - poor peer relationships, poor social interaction, more aggressive, conduct problems (Hoffman-Plotkin and Twentyman 1984; Rohrbeck and Twentyman 1986); delays in complex language (Eigsti et al. 2004); difficulties with discrimination of emotion expressions - bias for sad faces (Pollak et al. 2000); dysregulation emotion patterns (Maughan and Cicchetti 2002); helpless outlook, don't view others as a source of help (Macfie et al.	

		1999). Five to 6 years - insecure-avoidant attachment (Venet et al. 2007); poor peer relationships, rate self as angry, oppositional, others as sad/hurt (Macfie et al. 2001; Waldinger et al. 2001); low self-esteem (Toth et al. 1997); inclination to cheat and break rules (Koenig et al. 2004). Carer-child interaction features 0 to 12 months - low maternal sensitivity (Cicchetti et al. 2006); infants viewed as irritating and demanding (Cicchetti et al. 2006); use fewer commands and give less positive feedback to their infants (Christopoulos et al. 1988). One to 3 years - low attunement and lack competence (Fagan and Dore, 1993), withdrawn and uninvolved with their children (Crittenden and Bonvillian 1984; Crittenden and DiLalla 1988; DiLalla and Crittenden 1990; Mustillo et al. 2011); critical and/or ignore the child's cues for help (English et al. 2005; Koenig et al. 2000; Pianta et al. 1989). Three to 6 years - less affectionate (Pollitt et al. 1975; Toth et al. 1997); least number of utterances with their child (Eigsti and Cicchetti 2004); least likely to relieve distress in their child (Macfie et al. 1999). The review concludes that 'these features should alert social and health care professionals to children who warrant detailed evaluation and family intervention' (p772).	

16. Nolin P, Ethier L (2007) Using neuropsychological profiles to classify neglected children with or without physical abuse. Child Abuse and Neglect 31: 631–43

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: Aim of the study is to 1) investigate whether cognitive functions can differentiate neglected children with or without physical abuse compared to	**Participants:** Children and young people. Children and young people aged 6 to 12 currently receiving child protection services due to maltreated and a matched non-maltreated control group. **Sample characteristics:**	The study found no significant differences between the groups on the Comprehension of Instructions test of receptive language (F=1.31, p=0.173, ES=0.020). This measure also did not contribute to discriminant analysis between abused and non-abused children (no data reported).	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
comparison participants; 2) demonstrate detrimental impact of maltreatment on children. Study looks at a range of cognitive functions. Only data in relation to language development have been extracted here. **Methodology:** Cross-sectional study - Comparative observational study comparing cognitive functioning of children who have experienced neglect/neglect and physical abuse with non-abused children. **Country:** Canada. **Source of funding:** Other – Conseil Quebecois de la Recherche Sociale and Fonds Quebecois de Recherche sur la Societe et la Culture.	• Age - Group 1 (neglect with physical abuse): 9.3 (SD=2.0); Group 2 (neglect without physical abuse): 8.7 (SD=1.9); Group 3 (comparison) 8.8 (SD=1.8). • Sex - Group 1: 61% male; Group 2: 54% male; Group 3: 51% male. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Group 1: 79% annual income <$25,000 (Canadian dollars); Group 2: 86% annual income <$25,000; Group 3: 77% annual income <$25,000. • Type of abuse - Neglect and physical abuse (n=56); neglect without physical abuse (n=28). • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Group 1 - neglect with physical abuse (n=56). • Group 2 - neglect without physical abuse (n=28). • Group 3 - comparison (n=53). • Total sample size N=137.		**Overall validity rating:** - Study limitations include the fact that only receptive, and not productive, language abilities were assessed. The study also had a relatively small sample size, particularly for the neglect without physical abuse subgroup.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Recognition indicators measured: Language - Receptive language measured using the comprehension of instructions subtest of the French-Canadian form of the NEPSY (A Developmental NEuroPSYchological Assessment, Korkman et al. 1998). Authors report that subtest has good reliability and validity (Korkman et al. 2003), although reliability scores are not reported.		

17. Noll JG, Shenk CE, Yeh MT et al. (2010) Receptive language and educational attainment for sexually abused females. Pediatrics 126: e615–22

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: Aim of the study to '… test whether the experience of childhood sexual abuse is associated with long-term receptive language acquisition and educational attainment deficits for females …' (pe615). **Methodology:** Other - prospective cross-sequential design, following participants up until the age of 30. Only data for <18 have	**Participants:** Children and young people. Sexually abused females aged between 6 and 16 at time of recruitment and were followed up over a 19-year time frame. **Sample characteristics:** • Age – Mean age (SD) and range at each assessment date - Time 1 (1987–9) - whole sample 11 (SD=3), 6–16; abused group 11 (SD=3), 6–16; comparison group 11 (SD=3), 6–16. Time 2 (1988–91) whole sample 12 (SD=3), 7–18; abused group 12 (SD=3), 7–17; comparison group 12 (SD=3), 7–18. Time 3 (1990–2) whole sample 13 (SD=3), 8–20; abused group 13	Changes in scores over time were analysed using hierarchical linear modelling (HLM). The sample as a whole showed significant improvement in receptive language scores over time (linear slope coefficient significantly different from 0 ($t(1,245)=-4.62$, $p<0.001$). The improvement tended to level off in early twenties, as shown by quadratic coefficient significantly different from 0 ($t(1,244)=10.00$, $p<0.001$). There was no significant between the groups at age 6 ($t(1,207)=0.04$, $p=0.96$). However, there was a significant group x linear time interaction ($t(1,243)=2.68$, $p=0.008$), and group x quadratic time interaction ($t(1,254)=-2.41$, $p=0.01$).	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + No information provided about whether assessors were blind to participant group. Study did not appear to repeat measures of socioeconomic status

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
been extracted here. This study has been critically appraised using a prospective cohort tool. **Country:** USA, Washington.	(SD=3), 8–18; comparison group 13 (SD=3), 8–20. Time 4 (1996–8) whole sample 18(SD=4), 11–25; abused group 19 (SD=4), 11–25; comparison group 18 (SD=3), 11–23. Time 5 (1999–2001) whole sample 20 (SD=3), 13–26; abused group 21(SD=3), 13–26; comparison group 20 (SD=3), 13-26. Time 6 (2004–6) - whole sample 24 (SD=3), 18–30; abused group 25 (SD=4), 18–24; comparison group 24 (SD=3), 18–30. Number per developmental period - childhood (age 6–10) - whole sample 63; abused group 32; comparison group 31. Young/mid-adolescence (age 11–14) - whole sample 113; abused group 54; comparison group 59. Mid/late adolescence (age 15–19) - whole sample 106; abused group 52; comparison group 54. Young adulthood (19–25) – whole sample 145, abused group 65; comparison group 80. Adulthood (age 26–30) - whole sample 62; abused group 27; comparison group 35. • Sex - All female sample. • Ethnicity - Authors report % 'minority' which is defined as Black, Hispanic or Asian. Whole sample 46%; abused group 39%; Comparison group 51%. • Religion/belief - Not reported.	General linear modelling was used to test for differences in language ability between abused and comparison groups at different developmental stages. This showed a non-significant (using corrected significance criterion of p=0.007) difference in receptive language abilities in childhood (6–10 years) ($F(1,62)=0.31$, $p>0.007$) and young/mid-adolescence (11–14 years) ($F(1,112)=6.09$, $p>0.007$). Significant differences were found in mid/late adolescence (15–18 years) ($F(1,105)=9.38$, $p<0.007$), young adulthood (19–25 years) ($F(1,144)=7.68$, $p<0.007$) and adulthood (26–30 years) ($F(1, 161)=7.59$, $p<0.007$). NB. The reviewing team sought to calculate effect sizes for significant differences. However, this was not possible as sample sizes for abused and comparison groups were not provided.	and other relevant factors, which may have contributed to language development over time.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position – Socioeconomic status as defined by Hollingshead ratings: Mean (standard deviation), range. Whole sample 36 (SD=12), 11–44; Abused group 35 (SD=14), 10–47; Comparison group 37(SD=11), 12–43. • Type of abuse - Substantiated sexual abuse include genital contact and/or penetration. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Abused group n=84. • Comparison group n=89. • Whole sample n=173. NB. Sample size appears to refer to sample at Time 1 and sample sizes at subsequent time points will differ. **Recognition indicators measured:** Language - Receptive language measured using the Peabody Picture Vocabulary Test-Revised (PPVT-R) (Dunn and Dunn 1981). Picture-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	prompted vocabulary test in which participants shown array of pictures and prompted to identify pictures that best represents target vocabulary word. Authors state that this is reliable (no data given) and shows convergent validity with Wechsler IQ scores.		

18. Pears K, Fisher PA (2005) Developmental, cognitive, and neuropsychological functioning in preschool-aged foster children: Associations with prior maltreatment and placement history. Journal of Developmental and Behavioral Pediatrics 26: 112–22

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To examine '… a range of domains (e.g. physical growth, neuropsychological function, general cognitive function, language and executive function) in young children in foster care compared to a community sample of same-aged children from comparable socioeconomic status (SES) backgrounds' (p113). **Methodology:** Cross-sectional study. Observational comparative study, comparing maltreated and non-maltreated children.	**Participants:** Children and young people. Children and young people aged 3 to 6, 99 of whom were in foster care following maltreatment and 54 comparable non-maltreated children. **Sample characteristics:** • Age - Foster children: 4.38 (SD=0.79); Community comparison: 4.26 (SD=0.77). • Sex - Foster children: 51% boys; Community comparison: 52% boys. • Ethnicity - Foster children: 89% European American, 1% African American, 5% Latino, 5% Native American Community comparison: 82% European American, 7% African American, 7% Latino, 2% Native American, 2% Pacific Islander. • Religion/belief - Not reported. • Disability - Not reported.	There was a significant difference in language ability, with children in foster care showing significantly lower ability than a community comparison (t=-4.44, p=0.000). The reviewing team calculated an effect size from the data in the paper which was medium to large (ES=-0.78). Amongst maltreated children, there was a significant correlation between presence of neglect or emotional abuse and poorer language ability, with small to medium effect size (r=-0.22, p<0.05). There was a significant positive association between the number of maltreatment types children had experienced, and better language ability, with small to medium effect size (r=0.23, p<0.05). The authors hypothesise that this may be because children who have experienced more types of abuse come to the attention of authorities earlier and are so more likely to receive services.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Country: USA. **Source of funding:** Government – Grants from the US National Institute of Mental Health and Office for Research on Minority Health.	• Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Socioeconomic status of foster children not ascertained, except for 23 who had returned to their biological families, for whom median income was $10,000 to $14,999. Median income for community comparison $15,000 to $19,999. Authors report that this means that 'half the families were below or around the poverty level' (p114). Significant difference between highest levels of education attained: Foster care median = high school community comparison median = college or vocational school courses. • Type of abuse - Maltreatment history available for 94 of 99 foster children: 61% neglect, 17% sexual abuse, 14% physical abuse, 8% emotional abuse. • Looked after or adopted status – Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Foster children (n=99). • Community comparison (n=54). • Total sample (n=153).		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Recognition indicators measured: Language - Language measured using the language domain of the NEPSY (A Developmental NEuroPSYchological Assessment (Korkman et al. 1998) and the Preschool Language Scale 3rd Edition (Zimmerman et al. 1991).		

19. Prasad MR, Kramer LA, Ewing-Cobbs L (2005) Cognitive and neuroimaging findings in physically abused preschoolers. Archives of Disease in Childhood 90: 82–5

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To '… characterise the cognitive, motor and language skills of toddlers and preschoolers who had been physically abused and to obtain concurrent MRIs of the brain …' (p82). **Methodology:** Cross-sectional study. **Country:** USA – Houston.	**Participants:** Children and young people. Children aged 14 to 77 months who had been hospitalised due to abusive injuries and a matched comparison group. **Sample characteristics:** • Age - Age in months – Physically abused group - 35.25 (SD=19.80); comparison group - 29.74 (SD=19.89). • Sex - Physically abused group 10 males (53%); comparison group 8 males (42%). • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported.	The study found that children in the physically abused group had significantly lower scores on measures of receptive language than children in the community comparison group (F(1,36)=9.49, p=0.004) and expressive language (F(1,36)=13.68, p=0.0007). The reviewing team calculated effect sizes for these variables using mean and standard deviation data reported in the paper. For receptive language there was a large effect size of -1.00, and for expressive language, there was a large effect size of -1.23.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + Relatively small sample size. Consideration not given to generalisability of sample, given that some participants are children who have been hospitalised due to maltreatment, and so are suffering relatively severe abuse.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Socioeconomic position - Physically abused group - Hollingshead group I–III - 6 (32%), group IV–V 13 (68%); comparison group - Hollingshead group I–III - 6 (32%), group IV–V 13 (68%). • Type of abuse - Physical abuse (n=19). Burns - face (1), hands/arms (4), feet/legs (6), genitalia (3), body (1), fractures - fibula (1), tibia (1), humerus (1), clavicle (1), femur (5). Bruises/lacerations - hands/arms (6), feet/legs (3), face (7), genitalia (2), torso (7). Organ contusion/laceration - lungs (1), liver (1). • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Physically abused group (n=19). • Comparison group (n=19). • Total sample (n=38). **Recognition indicators measured** Language - Language measured using Sequenced Inventory of Communication Development (Hendrick et al. 1995); for children over 36 months Clinical Evaluation of Language Fundamentals (Preschool or Third Edition) (Semel et al. 1995). Reliability of instruments not reported.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Abuse determined by Child Protection Committee at each hospital and state protective regulatory agency.		

20. Rhodes AE et al. (2011) Sex differences in childhood sexual abuse and suicide related behaviors. Suicide & life-threatening behavior 41: 235–54

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The review aims to provide an update to earlier systematic reviews which aimed to determine the links between childhood sexual abuse and suicide-related behaviours. There is a particular focus on whether the strength of association differs in boys and girls. **Methodology:** Systematic review of empirical studies reporting unadjusted or adjusted results in relation to the association between sexual abuse and suicide-related behaviours. It is not clear whether the review protocol specified eligible study designs however the	**Participants:** Children and young people. Children and young people under the age of 18, or in grade 12 or lower. Studies with a clinical or child welfare based sample were excluded '… given that sexual abuse in these samples is formally disclosed and therefore may represent a different type of exposure than reported in the general population … Furthermore, the effects of disclosure, such as being separated from the parent, may modify the association' (p237). Childhood sexual abuse was defined as '… exposure to unwanted sexual acts or forced participation in sexual acts …' (p237). Suicide-related behaviours were defined (following Silverman et al., 2007) as '… self-harm (no suicidal intent), suicide-related behavior with undetermined intent, and suicide attempt' (p236). **Sample characteristics:**	**Statistical data** <u>Unadjusted associations between childhood sexual abuse and suicide attempt(s)</u> NB. Suicide was categorised dichotomously by the included studies. An odds ratio exceeding 1 denotes that the association was in the expected direction (i.e. that children who reported abuse were more likely to report suicide attempt/s). The review authors state that results with a confidence interval which does not include 1 are statistically significant. Eight studies provided unadjusted data on the association between sexual abuse and suicide attempts. Each of these found a significant association between childhood sexual abuse and suicide attempt(s) in both boys and girls. These studies also found that the magnitude of effect was greater in boys than in girls, and this difference was statistically significant in 7 of the 8 studies. The reported odds ratios for girls ranged between 2.2 and 11.2. The reported odds ratios for boys ranged between 1.8 and 30.8. Ackard and Neumark-Sztainer, 2003 (n=81, 247): Female: Girls who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 4.7 (4.5-5.0 95% CI). Male: Boys who reported abuse	**Overall assessment of internal validity:** + As the study did not include formal quality assessment of the included studies it is not possible to award a higher quality rating. **Overall assessment of external validity:** ++ **Overall validity rating:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
authors note that case reports, editorials, reviews and qualitative research were excluded and that the majority of included studies were cross-sectional observational studies. Ten of the included studies adjusted for potential confounders. **Country:** Range of countries. The included studies were conducted in: Ackard and Neumark-Sztainer, 2003 - USA. Anteghini et al., 2001 – Brazil. Bagley et al., 1995 – Canada. Bergen et al., 2003 – Australia. Borowsky et al., 1999 - USA. Choquet et al., 1997 – France. Edgardh and Ormstad, 2000 – Sweden. Eisenberg et al., 2007 - USA. Garnefski and Arends, 1998 – The Netherlands. Gold, 1996 - USA. Hawton et al., 2002 – United Kingdom - England. Howard and Wang, 2005 - USA. King et al., 2004 –	• Age - Percentages not reported – Ackard and Neumark-Sztainer, 2003 – Grades 9 and 12. Anteghini et al., 2001 – Unclear, reported as grades 8 and 10, and 13–17 years. Bagley et al., 1995 – Grades 7–12, ages 12–18. Bergen et al., 2003 – Year 9, mean age 14 years. Borowsky et al., 1999 – Grades 7-12. Choquet et al., 1997 – Grades 8-12, mean age 16.2. Edgardh and Ormstad, 2000 – 17 years. Eisenberg et al., 2007 - Grades 6, 9 and 12. Garnefski and Arends, 1998 – Unclear, reported as 11–23 years and 12–19 years. Gold, 1996 – Grade 10, mean age 15.5 years. Hawton et al., 2002 – 15 and 16 years. Howard and Wang, 2005 – Grades 9–12. King et al., 2004 – Grades 8 and 11, mean age 15.7 years. Martin et al., 2004 – Year 9, mean age 14 years. O'Connor et al., 2009 – years S4 or S5, ages 15-16 years. Olshen et al., 2007 – Grades 9–12, ages 14 and over. Rosenberg et al., 2005 – Grades 9–12, ages 13–18 years. • Sex - Ackard and Neumark-Sztainer, 2003 – Male 49.6%, female 50.4%. Anteghini et al., 2001 - Unclear. The authors' report the gender balance of a	were significantly more likely to report suicide attempt(s) than those who did not report abuse; 11.6 (10.7-12.9 95% CI). Female vs. male: Boys who reported abuse were significantly more likely than girls who reported abuse to report suicide attempt(s) – statistical data not presented. Anteghini et al., 2001 (n=2059): Female – Girls who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 3.4 (1.5-7.4 95% CI). Male – Boys who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 12.4 (5.3-29.0 95% CI). Female vs. male – Boys who reported abuse were significantly more likely than girls who reported abuse to report suicide attempt(s) – statistical data not presented. Bergen et al., 2003/Martin et al., 2004 (n=2485): Female - Girls who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 5.1 (2.5-10.4 95% CI) Male - Boys who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 30.8 (12.0-78.6 95% CI). Female vs. male: Boys who reported abuse were significantly more likely than girls who reported abuse to report suicide attempt(s) – statistical data not presented. Borowsky et al., 1999 (n=11, 666): Female: Girls who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 2.9 (2.5-3.4 95% CI) Male - Boys who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 4.8 (3.4-6.6 95% CI). Female vs. male: Boys who reported abuse were significantly more likely than girls who reported abuse to report suicide attempt(s) – statistical data not presented. Choquet et	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
South Africa. Martin et al., 2004 – Australia. O'Connor et al., 2009 – United Kingdom – Scotland. Olshen et al., 2007 – USA. Rosenberg et al., 2005 - USA. The review was conducted by authors based in Canada. **Source of funding:** • Government – Canadian Institutes of Health Research. • Other – Child Welfare League of Canada; Ontario Association for Children's Aid Societies; Ontario Centre of Excellence for Child and Youth Mental Health; The Injury and Child Maltreatment Section, Health Surveillance and Epidemiology Division, Public Health Agency of Canada, Centre of Excellence for Child Welfare.	subsample (male 44.6%, female 55.3%) which it is assumed are the subsample to whom the findings relate, however this is not made clear by the review authors. Bagley et al., 1995 – Male 51.5%, female 48.5%. Bergen et al., 2003 – Male 55.5%; female 44.5%. Borowsky et al., 1999 – Male 47.9%, female 52.1%. Choquet et al., 1997 - Male 48.7%, female 51.3%. Edgardh and Ormstad, 2000 - Male 41.9%, female 58.1%. Eisenberg et al., 2007 - Male 49.5%, female 50.5%. Garnefski and Arends, 1998 - Male 49.7%, female 50.1%. Gold, 1996 - Male 48.9%, female 51.1%. Hawton et al., 2002 - Male 52.9%, female 46.7%. Howard and Wang, 2005 – Not reported. King et al., 2004 - Male 43.5%, female 56.5%. Martin et al., 2004 - Male 55.5%, female 44.5%. O'Connor et al., 2009 - Male 46.6%, female 53.4%. Olshen et al., 2007 - Male 49.0%, female 51.0%. Rosenberg et al., 2005 – Not reported. • Ethnicity - Not reported for any of the included studies with the exception of Borowsky et al. 1999	al., 1997 (n=183): Authors of study did not provide unadjusted results. Eisenberg et al., 2007 (n=131, 862): Female: Girls who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 3.7 (3.5-3.9 95% CI). Male: Boys who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 11.2 (10.2-12.2 95% CI). Female vs male: Boys who reported abuse were significantly more likely than girls who reported abuse to report suicide attempt(s) – statistical data not presented. Female - non-familial perpetrator - suicide attempt – Girls who reported abuse by a non-familial perpetrator were significantly more likely to report suicide attempt(s) than those who did not report any abuse; 3.5 (3.3-3.8 95% CI). Female - familial perpetrator - suicide attempt - Girls who reported abuse by a familial perpetrator were significantly more likely to report suicide attempt(s) than those who did not report any abuse; 2.9 (2.5-3.2 95% CI). Female – familial and non-familial perpetrators - suicide attempt - Girls who reported abuse by both a familial and a non-familial perpetrator were significantly more likely to report suicide attempt(s) than those who did not report any abuse; 5.3 (4.7-5.8 95% CI). Male - non-familial perpetrator - suicide attempt – Boys who reported abuse by a non-familial perpetrator were significantly more likely to report suicide attempt(s) than those who did not report any abuse; 8.0 (7.1-9.0 95% CI). Male - familial perpetrator - suicide attempt - Boys who reported abuse by a familial perpetrator were significantly more likely to report suicide attempt(s) than those who did not report any abuse; 8.2 (6.7-10.1 95% CI). Male – familial and non-familial perpetrators - suicide attempt - Boys who reported	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	which focuses on American Indian and Alaska native students in reservation schools. • Religion/belief - Not reported for any of the included studies although Bagley et al., 1995 included students from Catholic schools. • Disability - Not reported for any of the included studies. • Long term health condition - Not reported for any of the included studies. • Sexual orientation - Not reported for any of the included studies. • Socioeconomic position - Not reported for any of the included studies. • Type of abuse - Ackard and Neumark-Sztainer 2003 – Date rape (lifetime) – male 1.6%, female 2.2%; unwanted sexual touching by adult/older person who is a family member (lifetime) - male 0.5%, female 2.5%; unwanted sexual touching by adult/older person who is not a family member (lifetime) - male 1.7%, female 6.1%; multiple forms of sexual abuse (lifetime) - male 2.3%, female 3.8%. Anteghini et al., 2001 – Sexual abuse (lifetime) - male 3.0%, female 3.4%. Bagley et al., 1995 – Sexual abuse outside of school	abuse by both a familial and a non-familial perpetrator were significantly more likely to report suicide attempt(s) than those who did not report any abuse; 25.9 (22.1-30.4 95% CI). Female vs. male - non-familial perpetrator - suicide attempt - Boys who reported abuse by a non-familial perpetrator were significantly more likely than girls who reported abuse by a non-familial perpetrator to report suicide attempt(s) – statistical data not presented. Female vs male – familial perpetrator - suicide attempt - Boys who reported abuse by a familial perpetrator were significantly more likely than girls who reported abuse by a non-familial perpetrator to report suicide attempt(s) – statistical data not presented. Female vs. male - familial and non-familial perpetrators - Boys who reported abuse by both a familial and a non-familial perpetrator were significantly more likely than girls who reported abuse by both a familial and a non-familial perpetrator to report suicide attempt(s) – statistical data not presented. Garnefski and Arends, 1998 (n=1490): Authors of study did not provide unadjusted results. Gold, 1996 (n=1335): Authors of study did not provide unadjusted results. Howard and Wang, 2005 (n=13, 601): Female: Girls who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 3.2 (2.6-4.1 95% CI). Male: Boys who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 4.5 (3.3-6.1 95% CI). Female vs. male: Boys who reported abuse were more likely than girls who reported abuse to report suicide attempt(s); however this difference was not statistically significant – statistical data not presented. King et al., 2004 (n=939): Authors of study did not provide unadjusted results.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	settings (lifetime) - male 9.8%, female 23.6% Bergen et al., 2003 - Sexual abuse (lifetime) – male 2.0%; female 5.4%. Borowsky et al., 1999 - Sexual abuse by a member of the family or anyone else (lifetime) - male 3.0%, female 16.8%. Choquet et al., 1997 – Sexual intercourse (lifetime) - male 0.6%, female 0.9%. Edgardh and Ormstad, 2000 – Sexual abuse by adults or young person at least 5 years older (lifetime) - male 3.1%, female 11.2%. Eisenberg et al., 2007 - Unwanted sexual touching by adult/older person who is a family member (lifetime) - male 0.7%, female 1.9%; unwanted sexual touching by adult/older person who is not a family member (lifetime) - male 2.2%, female 6.0%; unwanted sexual touching by adult/older person (familial and non-familial, lifetime) - male 1.1%, female 1.8%. Garnefski and Arends, 1998 - Sexual abuse (lifetime) – Overall 5.9% (not clear how this relates to subsample). Gold, 1996 - Sexual abuse (lifetime) – overall 10.9%. Hawton et al., 2002 - Sexual abuse (lifetime) – male 2.3%; female 6.3%. Howard and Wang, 2005 - Sexual intercourse	Olshen et al., 2007 (n=8080): Female: Girls who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 2.2 (1.4-3.4 95% CI). Male: Boys who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 9.1 (5.0-16.5 95% CI). Female vs male: Boys who reported abuse were significantly more likely than girls who reported abuse to report suicide attempt(s) – statistical data not presented. Rosenberg et al., 2005 (n=16,644): Forced sexual intercourse and single suicide attempt – Female: Girls who reported forced sexual intercourse were significantly more likely to report a single suicide attempt than those who did not report forced sexual intercourse; 2.2 (1.8-2.7 95% CI). Male: Boys who reported forced sexual intercourse were significantly more likely to report a single suicide attempt than those who did not report forced sexual intercourse; 1.8 (1.3-2.3 95% CI). Female vs. male: Not measured/reported. Rosenberg et al., 2005 (n=16,644): Forced sexual intercourse and multiple suicide attempts – Female: Girls who reported forced sexual intercourse were significantly more likely to report multiple suicide attempts than those who did not report forced sexual intercourse; 11.2 (8.4-14.9 95% CI). Male: Boys who reported forced sexual intercourse were significantly more likely to report multiple suicide attempts than those who did not report forced sexual intercourse; 5.0 (4.0-6.2 95% CI). Female vs male: Not measured/reported. <u>Adjusted associations between childhood sexual abuse and suicide attempt(s)</u> NB. Suicide was categorised dichotomously by the included studies. An odds ratio exceeding 1 denotes	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	(lifetime) – male 5.1%; female 10.2%. King et al., 2004 – Attempted sexual intercourse – male 2.0%; female 13.3%. Sexual intercourse – male 5.0%; female 6.0%. Martin et al., 2004 - Sexual abuse (lifetime) – male 2.0%; female 5.4%. O'Connor et al., 2009 - Sexual abuse (lifetime) – male 2.3%; female 6.3%. Olshen et al., 2007 - Sexual intercourse (lifetime) – male 5.4%; female 9.6%. Rosenberg et al., 2005 - Sexual intercourse (lifetime) – Male 6.5%; female 10.2%. • Looked after or adopted status - Not reported for any of the included studies, although the Guideline Committee may wish to note that the review excluded studies using a child welfare based sample. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported for any of the included studies. **Sample size:** The total sample size of the included studies is not reported by the review authors. The number of participants who were sampled in each study ranged between 183 to over 130,000.	that the association was in the expected direction (i.e. that children who reported abuse were more likely to report suicide attempt/s). Results with a confidence interval which does not include 1 are statistically significant. Ten studies reported in 11 papers provided adjusted results for the association between sexual abuse and suicide attempts (Anteghini et al. 2001; Bergen et al. 2003; Choquet et al. 1997; Eisenberg et al. 2007; Garnefski and Arends 1998; Gold 1996; Howard and Wang 2005; King et al. 2004; Martin et al. 2004; Olshen et al. 2007; Borowsky et al. 1999), although not all of these reported results in full. Studies adjusted for a range of factors hypothesised to mediate the CSA-suicide association, including ethnicity, family living arrangements, drug use, self-image, being bullied, uncertainty over sexual orientation etc. Each of these studies found an association between childhood sexual abuse and suicide attempt(s) in girls; however this association was only found to be significant by five studies. All 10 of the studies also found an association between childhood sexual abuse and suicide attempt(s) in boys; however this association was only found to be significant in 9 studies. For 6 studies reporting both unadjusted and adjusted results, in four the adjusted association remained statistically significant in boys but not girls (Anteghini et al. 2001; Howard and Wang 2005; Martin et al. 2004; Olshen et al. 2007). In the remaining two (Borowsky et al. 1999; Eisenberg et al. 2007), the associations remained significant, with the magnitude of the association greater for boys than girls. The reported adjusted odds ratios for girls ranged between 1.1 (95% CI 0.8 to 1.7) and 6.8 (95% CI 4.5 to 10.2 95% CI). The reported odds ratios for boys ranged between 1.9 (95% CI 1.1 to 3.2) and 27.8 (95% CI 9.8	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	The review included 16 studies which focused on the association between sexual abuse and suicide-related behaviours. These are reported in 17 different papers (NB. Bergen et al., 2003 and Martin et al., 2004 both report on the same sample). The review authors report both of these papers as Martin et al. 2004, however the NCCSC references use the correct citations for clarity.) **Recognition indicators measured:** Suicidal thoughts/behaviour All data in relation to abuse and suicide-related behaviour is self-reported through respondent's answers to questionnaires. The majority were anonymous and were based on lifetime recall. The review does not provide many details in relation the questions used except to note that they were generally brief and that Hawton et al., 2002 used open ended questions that were rated by three '... independent reviewers with specific criteria' (p243, no further details provided.) Ackard and Neumark-Sztainer, 2003 – Self-report - Date rape (lifetime); unwanted sexual touching by adult/older person who	to 78.9). Ackard and Neumark-Sztainer, 2003 (n=81, 247): Adjusted results not reported by review authors as '… suicidal thoughts were examined together with attempts …' (p245) Anteghini et al., 2001 (n=2059): Female: Girls who reported abuse were more likely to report suicide attempt(s) than those who did not report abuse; however this association was not statistically significant; statistical data not presented by authors of study. Male: Boys who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 8.2, p<.001. Female vs. male: Not measured/reported. Bergen et al., 2003/Martin et al., 2004 (n=2485): Female: Girls who reported abuse were more likely to report suicide attempt(s) than those who did not report abuse; however this association was not statistically significant; statistical data not presented by authors of study. Male: Boys who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 15.0 (4.7-47.9 95% CI). Female vs. male: Not measured/reported. Borowsky et al., 1999 (n=11, 666): Female: Girls who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 1.5 (1.2-1.8 95% CI). Male: Boys who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 2.2 (1.4-3.4 95% CI). Female vs. male: Boys who reported abuse were more likely than girls who reported abuse to report suicide attempt(s) but this difference was not statistically significant – statistical data not presented. Choquet et al., 1997 (n=183) (Only provides adjusted results, and does not include odds ratios): Female – Girls who reported abuse were more likely to report suicide attempt(s) than those who did	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	is a family member (lifetime); unwanted sexual touching by adult/older person who is not a family member (lifetime); multiple forms of sexual abuse (lifetime). Self-report - Suicide attempt(s) (lifetime) – male 7.2%; female 15.1%. Anteghini et al., 2001 – Self-report - Sexual abuse (lifetime). Self-report – Suicide attempt(s) (lifetime) – male 7.6%; female 10.3%. Bagley et al., 1995 – Self-report - Sexual abuse outside of school settings (lifetime). Self-report – Deliberate attempts to hurt or kill self, often (lifetime) – male 2.1%; female 2.3%. Bergen et al., 2003 - Self-report - Sexual abuse (lifetime). Self-report – Suicide attempt(s) (lifetime) - male 4.5%; female 7.3%. Deliberate self-injury (lifetime) – male 17.5%, female 19.3%. NB This study is based on the same sample as that reported in Martin et al. 2004, both of which report different measures of suicide-related behaviours, however it is not clear from the way in which these papers are reported by the review whether these relate to separate subsamples. They have therefore been combined here. Borowsky et al., 1999 - Self-report - Sexual abuse by a member of the family or anyone else (lifetime).	not report abuse; however this association was not statistically significant – statistical data not presented; prevalence of suicide attempt(s) was 22% amongst those who reported rape and 2% amongst controls ($\chi 2$ analysis showed that this association was not significant). Male – Boys who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; prevalence of suicide attempt(s) was 52% amongst those who reported rape and 12% amongst controls (Fischer 2-tail $p<.001$). Female vs. male – Not measured/reported. Eisenberg et al., 2007 (n=131, 862): Female - non-familial perpetrator - suicide attempt - Girls who reported abuse by a non-familial perpetrator were significantly more likely to report suicide attempt(s) than those who did not report any abuse; 3.5 (3.2-3.7 95% CI). Female - familial perpetrator - suicide attempt - Girls who reported abuse by a familial perpetrator were significantly more likely to report suicide attempt(s) than those who did not report any abuse; 2.5 (2.2-2.9 95% CI). Female – familial and non-familial perpetrators - suicide attempt - Girls who reported abuse by both a familial and a non-familial perpetrator were significantly more likely to report suicide attempt(s) than those who did not report any abuse; 5.6 (4.9-6.4 95% CI). Male - non-familial perpetrator - suicide attempt - Boys who reported abuse by a non-familial perpetrator were significantly more likely to report suicide attempt(s) than those who did not report any abuse; 4.9 (4.3-5.6 95% CI). Male - familial perpetrator - suicide attempt - Boys who reported abuse by a familial perpetrator were significantly more likely to report suicide attempt(s) than those who did not report any abuse; 5.0 (4.0-6.3 95% CI). Male – familial and non-familial perpetrators - suicide attempt - Boys	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Self-report – Suicide attempt(s) (lifetime) – male 11.8%; female 21.8%. Choquet et al., 1997 – Self-report - Sexual intercourse (lifetime). Self-report - Suicide attempt(s) (lifetime). Prevalence not reported. Edgardh and Ormstad, 2000 – Self-report - Sexual abuse by adults or young person at least five years older (lifetime). Self-report - Suicide attempt(s) or other act (s) of self-harm (lifetime) – male 5.9%; female 11.4%. Eisenberg et al. 2007 - Self-report - Unwanted sexual touching by adult/older person who is a family member (lifetime); unwanted sexual touching by adult/older person who is not a family member (lifetime); unwanted sexual touching by adult/older person (familial and non-familial, lifetime). Self-report - Suicide attempt(s) (lifetime) – male 7.4%; female 11.9%. Garnefski and Arends, 1998 - Self-report - Sexual abuse (lifetime). Self-report – Serious suicide attempt(s) (lifetime). Prevalence not reported. Gold, 1996 - Self-report - Sexual abuse (lifetime). Self-report - Suicide attempt (lifetime) – male 9.4%; female 24.5%. Hawton et al., 2002 - Self-report - Sexual abuse (lifetime). Self-report – Deliberate self-harm (past year) – male 3.2%;	who reported abuse by both a familial and a non-familial perpetrator were significantly more likely to report suicide attempt(s) than those who did not report any abuse; 10.8 (8.9-13.1 95% CI). Female vs. male - non-familial perpetrator - suicide attempt - Boys who reported abuse by a non-familial perpetrator were significantly more likely than girls who reported abuse by a non-familial perpetrator to report suicide attempt(s) – statistical data not presented. Female vs male – familial perpetrator - suicide attempt - Boys who reported abuse by a familial perpetrator were significantly more likely than girls who reported abuse by a non-familial perpetrator to report suicide attempt(s) – statistical data not presented. Female vs male - familial and non-familial perpetrators - Boys who reported abuse by both a familial and a non-familial perpetrator were significantly more likely than girls who reported abuse by both a familial and a non-familial perpetrator to report suicide attempt(s) – statistical data not presented. Garnefski and Arends, 1998 (n=1490) (Only provides adjusted results): Female – Girls who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 6.8 (4.5-10.2 95% CI). Male – Boys who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 27.8 (9.8-78.9 95% CI). Female vs. male - Boys who reported abuse were significantly more likely than girls who reported abuse to report suicide attempt(s) – statistical data not presented. Gold, 1996 (n=1335) (Only provides adjusted results): Female – Girls who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 2.7 (1.6-4.7 95% CI). Male – Boys who reported abuse were more likely to report suicide	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	female 11.2%. Howard and Wang, 2005 - Self-report - Sexual intercourse (lifetime). Self-report - Suicide attempt(s) (past year). Prevalence not reported. King et al., 2004 – Self-report - Attempted sexual intercourse. Sexual intercourse. Self-report - Suicide attempt(s) (past year) – overall 9.8 %. Martin et al., 2004 - Self-report - Sexual abuse (lifetime). Self-report – Suicide attempt(s) (lifetime) - male 4.5%; female 7.3%. Deliberate self-injury (lifetime) – male 17.5%, female 19.3%. NB This study is based on the same sample as that reported in Bergen et al., 2003, both of which report different measures of suicide-related behaviours, however it is not clear from the way in which these papers are reported by the review whether these relate to separate subsamples. They have therefor been combined here. O'Connor et al., 2009 - Self-report - Sexual abuse (lifetime). Self-report – Self-harm (past year) male 5.1%; female 13.6%. Self-harm (lifetime) – male 6.9%; female 19.9%. Olshen et al., 2007 - Self-report - Sexual intercourse (lifetime). Self-report – Suicide attempt(s) (past year) – male 7.2%; female 11.7%. Rosenberg et al., 2005 - Self-report -	attempt(s) than those who did not report abuse; however this association was not statistically significant; 3.2 (0.90-11.1 95% CI). Female vs. male - Not measured/reported. Howard and Wang, 2005 (N=13, 601): Female: Girls who reported abuse were more likely to report suicide attempt(s) than those who did not report abuse; however this association was not statistically significant; 1.1 (0.8-1.7 95% CI). Male: Boys who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 1.9 (1.1-3.2 95% CI). Female vs. male - Not measured/reported. King et al., 2004 (n=939) (Only provides adjusted results) - Female – Girls who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 3.1 (1.5-6.4 95% CI). NB The review authors report this as non-significant. Male – Boys who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 6.8 (1.3-36.4 95% CI). Female vs. male - Boys who reported abuse were significantly more likely than girls who reported abuse to report suicide attempt(s) – statistical data not presented. Olshen et al., 2007 (n=8080): Female: Girls who reported abuse were more likely to report suicide attempt(s) than those who did not report abuse; however this association was not statistically significant - statistical data not presented by authors of study. Male: Boys who reported abuse were significantly more likely to report suicide attempt(s) than those who did not report abuse; 3.9 (2.1-7.1 95% CI). Female vs. male: Not measured/reported. <u>Unadjusted associations between childhood sexual abuse and suicide-related behaviours</u>	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Sexual intercourse (lifetime). Self-report – One suicide attempt (past year) – male 10.3%; female 10.2%. Two or more suicide attempts (past year) – male 3.2%; female 5.8%. NB. It should be noted that Choquet et al., 1997, Howard and Wang, 2005, King et al., 2004, Olshen et al., 2007, and Rosenberg et al., 2005 all defined childhood sexual abuse as involving intercourse. (No further details provided).	Five studies reported across 6 papers examined the association between sexual abuse and suicide-related phenomena (e.g. self-harm) where the intent was unknown (Bagley et al. 1995; Bergen et al. 2003/ Martin et al. 2004; Edgardh and Ormstad 2000; Hawton et al. 2002; O'Connor et al. 2009). For the unadjusted data, all studies found a statistically significant association between abuse and suicide-related behaviours in both boys and girls, with reported odds ratios for girls ranging from 3.3 (95% CI 1.8 to 5.5) to 4.1 (95% CI 3.0 to 5.6) and odds ratios for boys ranging from 2.9 (95% CI 2.9 to 19.2) to 10.3 (95% CI 4.0 to 26.0). After controlling for variables such as depression, family functioning and drug use the four studies reporting adjusted results (Bergen et al. 2003/ Martin et al. 2004; Edgardh and Ormstad 2000; Hawton et al. 2002; O'Connor et al. 2009) found that none of the associations between abuse and suicide-related behaviours in girls was statistically significant, and only one study found a statistically association in boys. No adjusted odds ratios for girls were reported, 1 adjusted odds ratio (for significant result) for boys was reported – 4.3 (95% CI 1.5 to 12.6). Bagley et al., 1995 (n=2112): Female: Girls who reported abuse were significantly more likely to report suicide-related behaviour(s) than those who did not report abuse; statistical data not reported by review authors. Male: Boys who reported abuse were significantly more likely to report suicide-related behaviour(s) than those who did not report abuse; statistical data not reported by review authors. Female vs. male: Non-significant difference in magnitude of effect between boys and girls – statistical data not presented. Bergen et al.; 2003/Martin et al., 2004 (n=2485): Female: Girls who reported abuse were significantly more likely to report	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		suicide-related behaviour(s) than those who did not report abuse; 3.3 (1.8-5.9 95% CI). Male: Boys who reported abuse were significantly more likely to report suicide-related behaviour(s) than those who did not report abuse; 10.3 (4.0-26.0 95% CI). Female vs male: Not measured/reported. Edgardh and Ormstad, 2000 (n=1943): Female: Girls who reported abuse were significantly more likely to report suicide-related behaviour(s) than those who did not report abuse; statistical data not reported by review authors. Male: Boys who reported abuse were significantly more likely to report suicide-related behaviour(s) than those who did not report abuse; statistical data not reported by review authors. Female vs. male: Non-significant difference in magnitude of effect between boys and girls – statistical data not presented. Hawton et al., 2002 (n=6020): Female: Girls who reported abuse were significantly more likely to report suicide-related behaviour(s) than those who did not report abuse; 4.1 (3.0-5.6 95% CI). Male: Boys who reported abuse were significantly more likely to report suicide-related behaviour(s) than those who did not report abuse; 3.5 (1.6-7.9 95% CI). Female vs. male: Non-significant difference in magnitude of effect association between boys and girls – statistical data not presented. O'Connor et al., 2009 (n=2008): Female: Girls who reported abuse were significantly more likely to report suicide-related behaviour(s) than those who did not report abuse; 4.0 (2.4-6.7 95% CI; lifetime prevalence of childhood sexual abuse and suicide-related behaviours). Male: Boys who reported abuse were significantly more likely to report suicide-related behaviour(s) than those who did not report abuse; 2.9 (2.9-19.2 95% CI; lifetime prevalence of childhood sexual abuse and suicide-related behaviours). Female vs	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		male: Non-significant difference in magnitude of effect association between boys and girls – statistical data not presented. <u>Adjusted associations between childhood sexual abuse and suicide-related behaviours</u> Three studies reported adjusted results showing the association between childhood sexual abuse and suicide-related behaviours. Each of these found an association between abuse and suicide-related behaviours in girls, however none of these were significant. The 3 studies also found an association between abuse and suicide-related behaviours in boys, however this was only found to be significant in one study. None of the studies measured or reported on the difference in magnitude of effect between boys and girls. Bagley et al., 1995 (n=2112): Adjusted results not reported by individual study or review authors. Bergen et al., 2003/Martin et al., 2004 (n=2485): Female: Girls who reported abuse were more likely to report suicide-related behaviour(s) than those who did not report abuse; however this association was not statistically significant; statistical data not presented. Male: Boys who reported abuse were significantly more likely to report suicide-related behaviour(s) than those who did not report abuse; 4.3 (1.5-12.6 95% CI). Female vs. male: Not reported/measured. Edgardh and Ormstad, 2000 (n=1943): Adjusted results not reported by individual study or review authors. Hawton et al., 2002 (n=6020): Adjusted results not reported by study authors. Female: Girls who reported abuse were more likely to report suicide-related behaviour(s) than those who did not report abuse; however this association was not statistically significant – results not presented by study authors. Male: Boys	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		who reported abuse were more likely to report suicide-related behaviour(s) than those who did not report abuse; however this association was not statistically significant – results not presented by study authors. Female vs male: Not reported/measured. O'Connor et al., 2009 (n=2008): Adjusted results not reported by study authors. Female: Girls who reported abuse were more likely to report suicide-related behaviour(s) than those who did not report abuse; however this association was not statistically significant – results not presented by study authors. Male: Boys who reported abuse were more likely to report suicide-related behaviour(s) than those who did not report abuse; however this association was not statistically significant – results not presented by study authors. Female vs. male: Not reported/measured. **Narrative findings** The review included 17 papers reporting 16 studies which examined sex differences in the relationship between sexual abuse and suicide-related behaviours including self-harm (no suicidal intent), suicide-related behaviour with undetermined intent and suicide attempt, amongst 12–18-year-olds. Included studies were conducted in the US (7 studies), UK (2 studies), Australia (2 studies), France (1 study), Canada (1 study), Brazil (1 study), South Africa (1 study) and Sweden (1 study). The age of participants ranged between 11 and 18. Eight studies provided unadjusted data on the association between sexual abuse and suicide attempts (Ackard and Newmark-Sztainer 2003; Anteghini et al. 2001; Eisenberg et al. 2007; Howard and Wang 2005; Martin et al. 2004; Olshen et al. 2007; Rosenberg et al. 2005; Wagman Borowsky	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		et al. 1999*). There was a positive, statistically significant association between sexual abuse and suicide attempts in all 8 studies. Odds ratios were higher for boys than girls in all studies except for 1 (Rosenberg et al. 2005). Unadjusted odds ratios for girls ranged from 2.2 to 5.1, and unadjusted odds ratios ranged from 4.5 to 30.8 for boys. Ten studies reported in 11 papers provided adjusted results for the association between sexual abuse and suicide attempts (Anteghini et al. 2001; Bergen et al. 2003; Borowsky et al. 1999; Choquet et al. 1997; Eisenberg et al. 2007; Garnefski and Arends 1998; Gold 1996; Howard and Wang 2005; King et al. 2004; Martin et al. 2004; Olshen et al. 2007), although not all of these reported results in full. Studies adjusted for a range of factors hypothesised to mediate the CSA-suicide association, including ethnicity, family living arrangements, drug use, self-image, being bullied, uncertainty over sexual orientation, etc. Each of these studies found an association between childhood sexual abuse and suicide attempt(s) in girls; however this association was only found to be significant by 5 studies. All 10 of the studies also found an association between childhood sexual abuse and suicide attempt(s) in boys; however this association was only found to be significant by 9 studies. For 6 studies reporting both unadjusted and adjusted results, in 4 the adjusted association remained statistically significant in boys but not girls (Anteghini et al. 2001; Howard and Wang 2005; Martin et al. 2004; Olshen et al. 2007). In the remaining 2 (Borowsky et al. 1999; Eisenberg et al. 2007), the associations remained significant, with the magnitude of the association greater for boys than girls. The reported adjusted odds ratios for girls ranged between 1.1 (95% CI 0.8 to 1.7) and 6.8 (95% CI 4.5 to 10.2	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		95% CI). The reported adjusted odds ratios for boys ranged between 1.9 (95% CI 1.1 to 3.2) and 27.8 (95% CI 9.8 to 78.9). Five studies reported across 6 papers examined the association between sexual abuse and suicide-related phenomena (e.g. self-harm) where the intent was unknown (Bagley et al. 1995; Bergen et al. 2003/ Martin et al. 2004; Edgardh and Ormstad 2000; Hawton et al. 2002; O'Connor et al. 2009). For the unadjusted data, all studies found a statistically significant association between abuse and suicide-related behaviours in both boys and girls, with reported odds ratios for girls ranging from 3.3 (95% CI 1.8 to 5.5) to 4.1 (95% CI 3.0 to 5.6) and odds ratios for boys ranging from 2.9 (95% CI 2.9 to 19.2) to 10.3 (95% CI 4.0 to 26.0). After controlling for variables such as depression, family functioning and drug use the 4 studies reporting adjusted results (Bergen et al. 2003/Martin et al. 2004; Edgardh and Ormstad 2000; Hawton et al. 2002; O'Connor et al. 2009) found that none of the associations between abuse and suicide-related behaviours in girls was statistically significant, and only 1 study found a statistically association in boys. No adjusted odds ratios for girls were reported, one adjusted odds ratio (for significant result) for boys was reported – 4.3 (95% CI 1.5 to 12.6).	

21. Spratt EG, Friedenberg S, LaRosa A et al. (2012) The effects of early neglect on cognitive, language, and behavioral functioning in childhood. Psychology 3: 175–82

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: Purpose of the study is to '… compare cognitive, language and behavioral functioning of	**Participants:** • Children and young people - Participants were children between the ages of three and ten divided in to three groups: 1) children with a history of	One-way analysis of covariance comparing scores across three groups.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
children with no history of neglect to children with early neglectful situations, specifically those who experience physical and emotional neglect from a caregiver or deprivation due to pre-adoptive placement in an international institution environment …' (p175). **Methodology:** Cross-sectional study. Observational comparative study. **Country:** USA – South Carolina. **Source of funding:** Government – The study was supported by grants from the USA National Institutes of health and mental health.	physical or emotional neglect as defined by the Barnett Child Maltreatment Classification Scheme, 2) children adopted from international institutions 3) children with no history of neglect. • Caregivers and families - Caregivers of participating children were also involved in the study. **Sample characteristics:** • Age - Control group 67 months (SD=21.4), US neglected group 64 months (SD=26.9), internationally adopted group 73 months (SD=12.7). • Sex - Control group 15 male, 15 female; US neglected group 8 male, 9 female; internationally adopted group: 9 male, 6 female. • Ethnicity - Control group White n=20, Black n=6, other n=2; US neglect group White n=12, Black n=2, other n=3; internationally adopted group White n=14, Black n=0, other n=1. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position – Mean annual household income - control group $109,019 (SD=54995); US neglected group $37,889 (SD=22031);	<u>Test of Early Language Development – receptive measure</u>: There was a significant difference between the groups on the TELD receptive measure (F=9.33, p<0.0001). Post hoc tests showed that the control group performed significantly better than the US neglected group (p=0.004) and the internationally adopted group (p=0.002). It is not clear whether there was a difference between the US neglect and internationally adopted groups. It was not possible to calculate effect sizes from the available data. <u>Test of Early Language Development – expressive measure</u> There was a significant difference between the groups on the TELD expressive measure (F=8.96, p=0.0001). Post hoc tests showed that the control group performed significantly better than the US neglected group (p=0.006) and the internationally adopted group (p=0.001). It is not clear whether there was a difference between the US neglect and internationally adopted groups. It was not possible to calculate effect sizes from the available data. <u>Test of Early Language Development - oral composite measure</u> There was a significant difference between the group on the TELD oral composite measure (F=10.69, p<0.0001). Post hoc tests showed that the control group performed significantly better than the US neglected group (p=0.002) and the internationally adopted group (p=0.001). It is not clear whether there was a difference between the US neglect and internationally adopted groups. It was not possible to calculate effect sizes from the available data.	**Overall validity rating:** + Relatively small sample size, although statistically significant results still obtained. Physical measurements of participants taken and unclear how these were used.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	internationally adopted group $120,466 (SD=68376). • Type of abuse - Of the 32 children from the US neglect and international adoption groups 18 (56.3%) were known to have experienced physical neglect, 6 (18.8%) experienced medical neglect, 7 (21.9%) experienced physical abuse, 1 (3%) experienced sexual abuse and 3 (9.4%) experienced emotional abuse. 7 (21.9%) witnessed domestic violence. The internationally adopted group had been living in a stable environment for an average of 51.6 months. The US neglect group had been living in a stable environment for an average of 27.5 months. • Looked after or adopted status - 15 participants had been adopted from international institutions. Of the children born in the US, all were now living in a stable environment with extended family. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Control group n=28. • US neglect group n=17. • Internationally adopted group n=15. **Recognition indicators measured:** Language - Language functioning		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	measured using - for children aged up to 6 years 11 months - the Test of Early Language Development (TELD) (Hresko et al. 1981). This measures both receptive and expressive language, and an overall oral language composite. For children aged 7 to 9 years the Test of Language Development (TOLD) (Hammill and Newcomer 2010). Examines nine sub-categories of oral language competency. For children aged over 9 years the TOLD-intermediate.		

22. Stith SM, Liu T, Davies LC (2009) Risk factors in child maltreatment: A meta-analytic review of the literature. Aggression and Violent Behavior 14: 13–29

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The study uses a meta-analytic design to determine the strength of the relationship between a range of risk factors and abuse or neglect. **Methodology:** Systematic review. A meta-analysis of 155 studies examining 39 risk factors. The categorisation of risk factors is guided by ecological theory (Bronfenbrenner 1979), and	**Participants:** • Children and young people. Included studies had to empirically examine the relationship between a risk factor and either child physical abuse or child neglect. A definition of 'child' is not given. Reviewing team contacted authors of the paper who confirmed that included studies had to refer to children under 18 years of age. Details of individual studies not given. • Caregivers and families. Included studies were those in which perpetrator of abuse was parent or carer. Details about individual studies not given.	Review report gives values of both Cohen's d and Pearson's r. Here we have reported r as the more typical measure of association between two variables. Note that the authors use an interpretation of r which classifies effect sizes for r as large if >0.30, medium from 0.20 to 0.30 and small if they are from 0.10 to 0.20. We were unable to find the reference for this interpretation, and so have used the more conservative conventional values of 0.1=small, 0.3=medium, 0.5=large. <u>A - Effect sizes for physical abuse</u> <u>1. Relevant to Q4 - Parent-child interaction/parental report of child behaviour</u>	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** - Search limited to 1 database, and keyword searching only rather than free text searching. Unclear why effect sizes using both d and r have been calculated,

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
focuses on four groups of 'microsystemic' risk factors: - Parent-child interactions/parental report of child behaviour (relates to Q4) - Parent characteristics independent of the child (relates to Q4) - Child characteristics, excluding parents (relates to Q3) - Family characteristics (relates to Q4). **Country:** Not reported. Countries in which included studies were conducted is not reported. **Source of funding:** Government. Funded by Cooperative State Research, Education and Extension Service, US Department of Agriculture, US Air Force.	**Sample characteristics:** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Study considers physical abuse and neglect only. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** Systematic reviews - number of studies -155 included studies. Not possible to calculate total number of participants from information given in the paper. **Recognition indicators measured:** Risk factors - Meta-analysis examines 39 risk factors, categorised in to four groups: 1. Parent-child interactions/parental report of child behaviour (relates to Q4) 2. Parent characteristics independent of the child (relates to Q4)	1.1 Parent perceives child as a problem Examined in 25 studies, n=3317. Significant positive association with physical abuse, with medium effect size (r=0.30, p<0.001). Significant heterogeneity found between studies (Q^w=57.68, p<0.001). 1.2 Unplanned pregnancy Examined in 2 studies, n=1490. Significant positive association with physical abuse, with small to medium effect size (r=0.28, p<0.001). No significant heterogeneity found between studies (Q^w =0.31, p=ns). 1.3 Parent-child relationships Examined in 32 studies, n=1624. Significant negative association with physical abuse, with small to medium effect size (r=-0.27, p<0.001). Significant heterogeneity found between studies (Q^w =117.68, p<0.001). 1.4 Parent use of corporal punishment Examined in 7 studies, n=703. Significant positive association with physical abuse, with small to medium effect size (r=0.26, p<0.001). No significant heterogeneity found between studies (Q^w =4.65, p=ns). 1.5 Parenting behaviours Examined in 25 studies, n=2956. Significant positive association with physical abuse, with small effect size (r=0.17, p<0.001). Significant heterogeneity found between studies (Q^w =130.85, p<0.001). 1.6 Stress over parenting Examined in 11 studies, n=2075. Significant positive association with physical abuse, with small effect size (r=0.07, p<0.001). Significant heterogeneity found between studies (Q^w =51.14, p<0.001). <u>2. Relevant to Q4 - Parent characteristics independent of the child</u>. Review examined evidence in relation to association between 19 parent characteristics and physical abuse.	and what the Pearson's correlations signify in relation to categorical variables (e.g. parent gender). A high number of analyses resulted in statistically significant values for the Qw measure of homogeneity, suggesting a high degree of heterogeneity across studies. The authors have nonetheless chosen to combine these studies, and heterogeneity is taken in to account in the limitations section only.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	3. Child characteristics, excluding parents (relates to Q3) 4. Family characteristics (relates to Q4). 1. Parent-child interactions/parental report of child behaviour (relates to Q4) • Parent perceives child as problem - measured using Child Behavior Checklist; Eyberg behaviour inventory; Washington symptom checklist; Revised Conners parent rating scale; revised symptom behaviour checklist; Becker bi-polar checklist; interview; coded observation; parent/caregiver involvement scale; parenting stress index; relatedness scales. • Unplanned pregnancy – no measures stated. • Parent-child relationships - measured using observation using strange situation procedure; emotion management interview; observation using behavioural observation scoring system; questionnaire; observation using measure of maternal stimulation; observation using standardised observation codes 3rd revision; observation using maternal coding device; observation using interactional language; observation using Patterson system; observation using maternal style scale; observation using	2.1 Anger/hyper reactivity Examined in 9 studies, n=345. Significant positive association with physical abuse, with medium effect size (r=0.34, p<0.001). Significant heterogeneity found between studies (Q^w =14.25, p<0.05). 2.2 Anxiety Examined in 8 studies, n=563. Significant positive association with physical abuse, with small to medium effect size (r=0.29, p<0.001). No significant heterogeneity found between studies (Q^w =4.39, p=ns). 2.3 Psychopathology Examined in 13 studies, n=8630. Significant positive association with physical abuse, with small to medium effect size (r=0.28, p<0.001). Significant heterogeneity found between studies (Q^w =62.21, p<0.001). 2.4 Depression Examined in 14 studies, n=8258. Significant positive association with physical abuse, with small to medium effect size (r=0.27, p<0.001). Significant heterogeneity found between studies (Q^w =46.18, p<0.001). 2.5 Self-esteem Examined in 11 studies, n=2485. Significant negative association with physical abuse, with small to medium effect size (r=-0.24, p<0.001). Significant heterogeneity found between studies (Q^w=32.92, p<0.001). 2.6 Poor relationship with own parents Examined in 11 studies, n=2997. Significant positive association with physical abuse, with small to medium effect size (r=0.22, p<0.001). Significant heterogeneity found between studies (Q^w=20.38, p<0.001). 2.7 Parent experienced childhood abuse Examined in 15 studies, n=3722. Significant positive association with physical abuse, with small to medium effect size (r=0.21, p<0.001). Significant heterogeneity found between studies (Q^w=78.55, p<0.001).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	dyadic parent-child interaction coding system; mother-child interaction task; Q sort block child rearing practices; Caldwell's home observation; coded observation using Barnard scales. • Parent use of corporal punishment - measured using parent daily report; interview. Parenting behaviours - measured using coded observation; emotion management interview; parental problem-solving measure; interview; parent opinion questionnaire; Michigan screening profile of parenting; questionnaire/survey; Vineland social maturity index; Developmental expectation questionnaire (Vineland Social Maturity Index); manifest rejection scale; observation using maternal style scale; observation using a role-play inventory; parent daily report; Hogan empathy test; adult-adolescent parent inventory; empathy scales. Stress over parenting - interview; parenting stress index, questionnaire; parenting sense of competence scale. 2. Parent characteristics independent of the child (relates to Q4) • Anger/hyper-reactivity - measured using Buss-Durkee hostility inventory; Michigan screening profile of parenting; state-trait anger expression inventory; questionnaire; Eyberg	2.8 Criminal behaviours Examined in 4 studies, n=1963. Significant positive association with physical abuse, with small to medium effect size (r=0.21, p<0.001). No significant heterogeneity found between studies (Q^w=0.66, p=ns). 2.9 Personal stress Examined in 22 studies, n=3114. Significant negative association with physical abuse, with small effect size (r=0,19, p<0.001). Significant heterogeneity found between studies (Q^w=50.74, p<0.001). 2.10 Social support Examined in 20 studies, n=10315. Significant positive association with physical abuse, with small effect size (r=-0.18, p<0.001). Significant heterogeneity found between studies (Q^w =65.32, p<0.001). 2.11 Alcohol abuse Examined in 3 studies, n=654. Significant positive association with physical abuse, with small effect size (r=0.17, p<0.001). Significant heterogeneity found between studies (Q^w =8.06, p<0.05). 2.12 Unemployment Examined in 8 studies, n=1263. Significant positive association with physical abuse, with small effect size (r=0.15, p<0.001). Significant heterogeneity found between studies (Q^w =29.57, p<0.001). 2.13 Parenting coping and problem-solving skills Examined in 4 studies, n=303. Significant positive association with physical abuse, with small effect size (r=-0.14, p<0.05). Significant heterogeneity found between studies (Q^w=7.54, p<0.05). 2.14 Single parenthood Examined in 22 studies, n=14223. Significant positive association with physical abuse, with small effect size (r=0.12, p<0.001). Significant heterogeneity found between studies (Q^w =108.23, p<0.001).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	child behaviour inventory; mood adjective checklist; laboratory test. • Anxiety - measured using Cattell's 16 personality factor questionnaire; state-trait anxiety inventory; multiple affect adjective checklist; international classification of disease; DSM-II diagnosis. Psychopathology - measured using DSM-I diagnosis; Tennessee self concept scale; Brief symptom inventory; state-trait anxiety inventory; Cornell medical index; Mini-mult; symptom checklist-90-revised; current and past psychopathology scales; diagnostic interview schedule; Rorschach; Depression self-esteem – California test of personality; index of self-esteem; interpersonal support evaluation list; Items from: parent child relations questionnaire; self-description and mate description form; Interview; parental attribution test; Questionnaire; Rosenberg self-esteem scale; Tennessee self-concept scale. • Poor relationship with own parents – Block child rearing practices report; childhood social network questionnaire; Interview; Michigan screening profile of parenting; questionnaire. • Parent experienced childhood abuse – Adaptations or items from conflict tactics scales; attachment and support systems questionnaire; Interview; Item from: survey on bringing	2.15 Parent age Examined in 31 studies, n=12146. Significant negative association with physical abuse, with small effect size (r=-0.10, p<0.001). Significant heterogeneity found between studies (Q^w=234.05, p<0.001). 2.16 Drug abuse Examined in 3 studies, n=654. Significant positive association with physical abuse, with small effect size (r=-.08, p<0.05). No significant heterogeneity found between studies (Q^w =2.18, p=ns). 2.17 Health problems Examined in 3 studies, n=286. Non-significant association with physical abuse, with small effect size (r=0.11, p=ns). No significant heterogeneity found between studies (Q^w=3.17, p=ns). 2.18 Parent gender Examined in 2 studies, n=7309. Significant positive association with physical abuse, with very small effect size (r=0.07, pp<0.001). No significant heterogeneity found between studies (Q^w =0.1, p=ns). 2.19 Approval of corporal punishment Examined in 5 studies, n=1674. Non-significant association with physical abuse, with very small effect size (r=0.05, p=ns). No significant heterogeneity found between studies (Q^w=5.65, p=ns). <u>3. Relevant to Q3 - Child characteristics, excluding parents.</u> Review examined evidence in relation to association between 7 child characteristics and physical abuse. 3.1 Child social competence Examined in 14 studies, n=1527. Significant negative association with physical abuse, with small to medium effect size (r=-0.26, pp<0.001). Significant heterogeneity found between studies (Q^w=27.46, p<0.05). 3.2 Child externalising behaviours Examined in 31 studies, n=2874. Significant positive association with	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	up children; psychosocial interview; questionnaire. Criminal behaviours – criminal records; interview; questionnaire. • Personal stress – checklist of stressful life events; hassles scale; health visitor questionnaire; interview. Items from: social readjustment rating scale; life experience survey; parenting stress Index; questionnaire; recent life changes questionnaire; schedule of recent experience family life form; social readjustment rating scale. Social support – attachment and support systems questionnaire; community relationships index; diagnostic interview schedule; family relationship index; index of social network strength; interpersonal support evaluation list; interview; maternal social support index; parenting stress index; questionnaire; social network map. alcohol abuse – diagnostic interview schedule; interview; questionnaire. • Unemployment – Not reported. • Parent coping and problem-solving skills – family environment scale; Michigan screening profile of parenting; problem solving inventory; social information form; social problem solving inventory (rev.); social support system inventory. • Single parenthood – Not reported. • Parent age – Not reported.	physical abuse, with small to medium effect size (r=0.23, pp<0.001). Significant heterogeneity found between studies (Q^w=135.69, p<0.001). 3.3 Child internalising behaviours Examined in 23 studies, n=2282. Significant positive association with physical abuse, with small effect size (r=0.15, pp<0.001). Significant heterogeneity found between studies (Q^w=50.62, p<0.001). 3.4 Child gender Examined in 13 studies, n=1702. Non-significant association with physical abuse, with very small effect size (r=0.04, p=ns). No significant heterogeneity found between studies (Q^w=6.1, p=ns). 3.5 Prenatal or neonatal problems Examined in 10 studies, n=1432. Non-significant association with physical abuse, with very small effect size (r=0.04, p=ns). No significant heterogeneity found between studies (Q^w=15.34, p=ns). 3.6 Child disability Examined in 4 studies, n=325. Non-significant association with physical abuse, with very small effect size (r=0.01, p=ns). No significant heterogeneity found between studies (Q^w=0.8, p=ns). 3.7 Child age Examined in 14 studies, n=3332. Non-significant association with physical abuse, with very small effect size (r=-0.02, p=ns). No significant heterogeneity found between studies (Q^w=12.63, p=ns). <u>4. Relevant to Q4 - Family characteristics</u> Review examined association between 7 family characteristics and physical abuse. 4.1 Family conflict Examined in 5 studies, n=170. Significant positive association with physical abuse,	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Drug abuse – Diagnostic interview schedule; interview; questionnaire. • Health problems – Cornell medical index; parenting Stress Index; questionnaire. • Parent gender – Not reported. • Approval of corporal punishment – Item from: general social science survey; interview; treatment evaluation inventory. 3. Child characteristics, excluding parents (relevant to Q3) – • Child social competence – California child Q-set; child behavior checklist; child behavior form; coded observation; developmental profile; instrumental and social competence scale; peer ratings; Rothenberg social sensitivity test; self-perception profile for children; teacher's report form of the child behaviour profile; teacher's report; teacher's rating scale of child's actual behaviour; vineland social maturity index. Child externalising behaviours – behavior problem checklist; California child q-set; checklist of child distress; checklist of child distress symptoms; child abuse potential inventory; child behavior checklist; child behavior form; coded observation; coded observation using: core conflictual relationship theme method; Conners teacher rating scale; diagnostic interview schedule	with medium effect size (r=0.39, p<0.001). Significant heterogeneity found between studies (Q^w=16.02, p<0.05). 4.2 Family cohesion Examined in 5 studies, n=183. Significant negative association with physical abuse, with medium effect size (r=-0.32, p<0.001). No significant heterogeneity found between studies (Q^w=3.02, p=ns). 4.3 Spousal violence Examined in 5 studies, n=773. Significant positive association with physical abuse, with small to medium effect size (r=0.22, p<0.001). No significant heterogeneity found between studies (Q^w=3.82, p=ns). 4.4 Marital satisfaction Examined in 8 studies, n=840. Significant negative association with physical abuse, with small effect size (r=-0.16, p<0.001). Significant heterogeneity found between studies (Q^w=14.45, p<0.05). 4.5 Family size Examined in 23 studies, n=11224. Significant positive association with physical abuse, with small effect size (r=0.15, p<0.001). Significant heterogeneity found between studies (Q^w=65.53, p<0.001). 4.6 Socio-economic status Examined in 16 studies, n=10321. Significant negative association with physical abuse, with small effect size (r=-0.14, pp<0.001). Significant heterogeneity found between studies (Q^w=41.45, p<0.001). 4.7 Non-biological parent in home Examined in 3 studies, n=302. Non-significant association with physical abuse, with very small effect size (r=-0.03, p=ns). No significant heterogeneity found between studies (Q^w=3.25, p=ns).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	form; interview; observation using: interactional language; parenting stress index/short form; Pittsburgh adjustment survey scale; revised behavior problem checklist; self and peer ratings: teachers report form of child behavior profile; self-perception profile for children; teacher's rating scale of child's actual behaviour; teacher's report form of the child behavior profile. • Child internalising behaviours – Behavior problem checklist; California child Q-set; checklist of child distress symptoms; child behavior checklist; child behavior problem checklist; child behavior profile; children's depression inventory; coded observation; revised Conners parent rating scale; Conners teacher rating scale; diagnostic interview schedule for children; Harter dimensions of depression profile for children; parenting stress index; Pittsburgh adjustment survey scale; preschool behavior questionnaire; revised behavior problem checklist; self and peer ratings; teacher's rating scale of child's actual behaviour; teacher's report form of the child behavior profile. • Child gender – Not reported. • Prenatal or neonatal problems – Diagnostic interview schedule for children; hospital records; Interview; questionnaire.	<u>B – Effect sizes for neglect</u> <u>5. Relevant to Q4 - Parent-child interaction/parental report of child behaviour</u> Review examined association between 4 characteristics of parent-child interaction/parental report of child behaviour and neglect. 5.1 Parent-child relationships Examined in 11 studies, n=400. Significant negative association with neglect, with medium to large effect size (r=-0.48, p<0.001). Significant heterogeneity found between studies (Q^w=58.16, p<0.001). 5.2 Parent perceives child as a problem Examined in 4 studies, n=87. Significant positive association with neglect, with medium to large effect size (r=0.41, p<0.001). No significant heterogeneity found between studies (Q^w=4.91, p=ns). 5.3 Parenting behaviours Examined in 8 studies, n=1016. Significant positive association with neglect, with small effect size (r=0.18, p<0.001). Significant heterogeneity found between studies (Q^w=35.24, p<0.001). 5.4 Stress over parenting Examined in 4 studies, n=307. Significant positive association with neglect, with small effect size (r=0.14, p<0.01). Significant heterogeneity found between studies (Q^w=40.82, p<0.001). <u>6. Relevant to Q4 - Parent characteristics independent of the child</u>. Review examined the association of 11 parent characteristics with neglect. 6.1 Personal stress Examined in 3 studies, n=386. Significant positive association with neglect, with medium effect size (r=0.38, p<0.001). Significant	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Child disability – Interview; questionnaire. Child age – Not reported. 4. Family characteristics – • Family conflict – coded observation; family concept inventory; family environment scale; Moos family environment scale. Family cohesion – Coded observation; family adaptability and cohesion evaluation scales-II; family environment scale; Moos family environment scale. • Spousal violence – Interview; items from: conflict tactics scales; questionnaire. Marital satisfaction – Interview; marital adjustment test; parenting stress index; questionnaire. • Family size – Not reported. • Socioeconomic status – Not reported. • Non-biological parent in home - Not reported.	heterogeneity found between studies (Q^w=24.84, p<0.001). 6.2 Anger/hyper-reactivity Examined in 3 studies, n=211. Significant positive association with neglect, with medium effect size (r=0.35, p<0.001). No significant heterogeneity found between studies (Q^w=1.98, p=ns). 6.3 Self-esteem Examined in 4 studies, n=184. Significant negative association with neglect, with medium effect size (r=0.33, p<0.001). No significant heterogeneity found between studies (Q^w=1.98, p=ns). 6.4 Psychopathology Examined in 8 studies, n=7652. Significant positive association with neglect, with small to medium effect size (r=0.25, p<0.001). Significant heterogeneity found between studies (Q^w=14.45, p<0.05). 6.5 Unemployment Examined in 4 studies, n=719. Significant positive association with neglect, with small to medium effect size (r=0.25, p<0.001). Significant heterogeneity found between studies (Q^w=8.69, p<0.05). 6.6 Depression Examined in 8 studies, n=8207. Significant positive association with neglect, with small to medium effect size (r=0.21, p<0.001). Significant heterogeneity found between studies (Q^w=15.93, p<0.05). 6.7 Poor relationship with own parents Examined in 7 studies, n=855. Significant positive association with neglect, with small effect size (r=0.19, p<0.001). Significant heterogeneity found between studies (Q^w=15.4, p<0.05). 6.8 Social support Examined in 13 studies, n=8582. Significant negative association with neglect, with small effect size (r=-0.16,	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		p<0.001). Significant heterogeneity found between studies (Q^w=55.11, p<0.001). 6.9 Parent experienced childhood abuse Examined in 6 studies, n=1417. Significant positive association with neglect, with small effect size (r=0.15, p<0.001). Significant heterogeneity found between studies (Q^w=25.21, p<0.001). 6.10 Parent age Examined in 9 studies, n=8120. Significant negative association with neglect, with very small effect size (r=-0.012, p<0.001). Significant heterogeneity found between studies (Q^w=19.94, p<0.05). 6.11 Single parenthood Examined in 9 studies, n=7751. Significant positive association with neglect, with very small effect size (r=0.08, p<0.001). No significant heterogeneity found between studies (Q^w=13.41, p=ns). <u>7. Relevant to Q3 – Child characteristics, excluding parents</u> Review examined the association between 5 child characteristics and neglect. 7.1 Child social competence Examined in 7 studies, n=584. Significant negative association with neglect, with medium effect size (r=-0.3, p<0.001). No significant heterogeneity found between studies (Q^w=10.03, p=ns). 7.2 Child externalising behaviours Examined in 17 studies, n=956. Significant positive association with neglect, with small to medium effect size (r=0.22, p<0.001). Significant heterogeneity found between studies (Q^w=35.37, p<0.01). 7.3 Child internalising behaviours Examined in 11 studies, n=922. Significant positive association with	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		neglect, with small effect size (r=0.11, p<0.001). Significant heterogeneity found between studies (Q^w=48.22, p<0.01). 7.4 Child gender Examined in 5 studies, n=961. Non-significant association with neglect, with very small effect size (r=0.01, p=ns). No significant heterogeneity found between studies (Q^w=0.28, p=ns). 7.5 Child age Examined in 8 studies, n=369. Non-significant association with neglect, with very small effect size (r=-0.01, p=ns). No significant heterogeneity found between studies (Q^w=13.49, p=ns). <u>8. Relevant to Q4 – Family factors</u> Review examined the association between 2 family factors and neglect. 8.1 Family size Examined in 12 studies, n=8546. Significant positive association with neglect, with small to medium effect size (r=0.26, p<0.001). Significant heterogeneity found between studies (Q^w=75.19, p<0.001). 8.2 Socioeconomic status Examined in 10 studies, n=7986. Significant negative association with neglect, with small effect size (r=-0.19, p<0.001). Significant heterogeneity found between studies (Q^w=36.29, p<0.001).	

23. Tonmyr L, Thornton T, Draca J et al. (2010 A review of childhood maltreatment and adolescent substance use relationship. Current Psychiatry Reviews 6(3): 223–34

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To '…identify the presence of an association between child maltreatment (ne-	**Participants:** Children and young people. Maltreated adolescents (school and community samples i.e. non-clinical; school samples do not include non-attendees). NB A	**Statistical data** Odds ratios (the review authors recorded 95% confidence intervals or significance levels when these were available).	**Overall assessment of internal validity:** + No formal quality appraisal.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
glect, witnessing domestic violence, physical, sexual and emotional maltreatment) and nicotine, alcohol and/or drug use/abuse among adolescents …' (p224). **Methodology:** Systematic review. **Country:** Range of countries. Acierno et al., 2000/Kilpatrick et al., 2000 – USA. Champion et al., 2004 – USA. Clark et al., 2004; 2005 – USA. Fergusson et al., 1997; 1996 – New Zealand. Behnken et al., 2010 – USA. Bergen et al., 2004 – Australia. Chandy et al., 1997 – USA. Choquet et al., 1997 – France. Edgardh and Ormstad, 2000 – Sweden. Erickson and Rapkin, 1991 – USA. Frederiksen et al., 2008 – Denmark. Garnefski and Arends, 1998 – Netherlands. Hamburger et al., 2008 – USA. Hernandez et al.,	number of the studies are linked to the same trial and appear to use the same sample. **Sample characteristics:** • Age - Acierno et al., 2000/Kilpatrick et al., 2000 – 12–17 years. Champion et al., 2004 – 16–20 years. Clark et al., 2004; 2005 – Not reported. Fergusson et al., 1997; 1996 – 0–18 years. Behnken et al., 2010 – 16 years. Bergen et al., 2004 – 13, 14, and 15 years. Chandy et al., 1997 – Mean age 15.3 years; grades 7-12. Choquet et al., 1997 – Mean age 16.2 years; grades 8–12. Edgardh and Ormstad, 2000 – 17 years. Erickson and Rapkin, 1991 – Grades 6–12. Frederiksen et al., 2008 – 15–16 years; grade 9. Garnefski and Arends, 1998 – 12–19 years. Hamburger et al., 2008 – grades 7, 9, 11, 12. Hernandez et al., 1992 - Hernandez et al., 1993 – Grade 9, and grades 9–12. Hibbard et al., 1988 – 11–17 years. Hibbard et al., 1990 – Grades 7–12. Howard et al., 2005 – Grades 9-12. Lau et al., 2003 – Grade 8. Logan et al., 2009 – Grade 7. Luster and Small, 1997 – Grades 7–12. Nagy et al., 1994 – Grades 8 and 10. Nelson et al., 1994 –	Neglect and alcohol use/abuse - The review identified 3 papers reporting on 2 studies which tested the association between neglect and alcohol use/abuse (Clark et al., 2004; Clark et al., 2005; and Shin et al., 2009). All 3 studies found an association between neglect and alcohol use/abuse, with respondents who reported neglect being more likely than those who did not report neglect to also report alcohol use abuse; however this association was only found to be significant in Clark et al., 2004 and Clark et al., 2005. The reported odds ratios of these studies ranged between 1.2 and 21.2. Clark et al., 2004: Respondents who reported neglect were significantly more likely than those who did not report neglect to report alcohol use/abuse; 3.2 odds ratio (95% CI 1.3 to 8.3). Clark et al., 2005: Respondents who reported neglect were significantly more likely than those who did not report neglect to report alcohol use/abuse; 21.2 odds ratio (95% CI 5.0 to 89.7). Shin et al., 2009: Respondents who reported neglect were more likely than those who did not report neglect to report alcohol use/abuse, however this association was not significant; 1.2 odds ratio (95% CI 1.0 to 1.5). NB Reported as significant by the review authors. Neglect and cigarette use – The review did not identify any papers which reported on the association between neglect and cigarette use. Neglect and 'drug' use/abuse – The review does not report on studies which measured the association between neglect and 'drug' use or abuse as it does for emotional abuse, physical abuse and sexual abuse.	**Overall assessment of external validity:** ++ **Overall validity rating:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
1992/1993 – USA. Hibbard et al., 1988 – USA. Hibbard et al., 1990 – USA. Howard et al., 2005 – USA. Lau et al., 2003 – Hong Kong. Logan et al., 2009 – USA. Luster and Small, 1997 – USA. Nagy et al., 1994 – USA. Nelson et al., 1994 – USA. Moran et al., 2004 – USA. Pedersen and Skrondal, 1996 – Norway. Perkins and Jones, 2004 – USA. Riggs et al., 1990 – USA. Shin et al., 2009 – USA. Simantov et al., 2000 – USA. Southwick-Bensley et al., 1999 – USA. Watts and Ellis, 1993 – USA. Yen et al., 2008 – Taiwan. The review was conducted by researchers based in Canada. **Source of funding:** Not reported.	Grades 9–12. Moran et al., 2004 – Grades 10–12. Pedersen and Skrondal, 1996 – Mean age 13.7 years. Perkins and Jones, 2004 – 12–17 years. Riggs et al., 1990 – Grades 9-12. Shin et al., 2009 – Grades 7–12. Simantov et al., 2000 – Grades 7–12. Southwick-Bensley et al., 1999 – Grades 8, 10, and 12. Watts and Ellis, 1993 – Grades 7–13. Yen et al., 2008 – 13–18 years. • Sex - Acierno et al., 2000/Kilpatrick et al., 2000 – Not reported. Champion et al., 2004 – Male and female. Clark et al., 2004; 2005 – Not reported. Fergusson et al., 1997; 1996 – Not reported. Behnken et al., 2010 – Female. Bergen et al., 2004 – Male and female. Chandy et al., 1997 – Male. Choquet et al., 1997 – Not reported. Edgardh and Ormstad, 2000 – Male and female. Erickson and Rapkin, 1991 – Not reported. Frederiksen et al., 2008 – Male and female. Garnefski and Arends, 1998 – Not reported. Hamburger et al., 2008 – Not reported. Hernandez et al., 1992 – Male and female. Hernandez et al., 1993 – Male. Hibbard et al., 1988 – Not reported. Hibbard et al., 1990 – Not reported. Howard et al., 2005 –	Neglect and use or abuse of other substances – The review does not report on studies which measured the association between neglect and use or abuse of other substances as it does for emotional abuse, physical abuse and sexual abuse (i.e. marijuana/hashish, cocaine/crack, methamphetamine, barbiturate, stimulants/uppers/speed, inhalants, hallucinogens, 'designer drugs', steroids, or medication/prescription). Witnessing domestic violence and alcohol use/abuse – The review identified two papers which reported on the association between witnessing domestic violence and alcohol use/abuse (Hamburger et al., 2008; and Simantov et al., 2000 – reported for both males and females). Both studies found an association between witnessing domestic violence and alcohol use/abuse, with respondents who reported domestic violence being more likely to report alcohol use/abuse than those who did not report witnessing domestic violence (for both females and males as reported in Simantov et al., 2000). Hamburger et al., 2008 found that the association was significant; whilst Simantov et al., 2000 found that the association was significant in females but not in males. The reported odds ratios of these studies ranged between 1.4 and 1.9. Hamburger et al., 2008: Respondents who reported domestic violence were significantly more likely than those who did not report witnessing domestic violence to report alcohol use abuse; 1.9 (95% CI 1.6 to 2.2). Simantov et al., 2000 (males): Male respondents who reported witnessing domestic violence were more likely than those who did not report witnessing domestic violence to report alcohol use/abuse, but this association was not significant; 1.4 odds ratio (95% CI 0.9 to 2.0).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Male and female. Lau et al., 2003 – Not reported. Logan et al., 2009 – Male and female. Luster and Small, 1997 – Not reported. Nagy et al., 1994 – Not reported. Nelson et al., 1994 – Not reported. Moran et al., 2004 – Not reported. Pedersen and Skrondal, 1996 – Male and female. Perkins and Jones, 2004 – Not reported. Riggs et al., 1990 – Not reported. Shin et al., 2009 – Not reported. Simantov et al., 2000 – Not reported. Southwick-Bensley et al., 1999 – Not reported. Watts and Ellis, 1993 – Not reported. Yen et al., 2008 – Not reported. • Ethnicity - Not reported for any of the included studies. • Religion/belief - Not reported for any of the included studies. • Disability - Not reported for any of the included studies. • Long term health condition - Not reported for any of the included studies. • Sexual orientation - Not reported for any of the included studies. • Socioeconomic position - Not reported for any of the included studies.	Simantov et al., 2000 (females): Female respondents who reported domestic violence were significantly more likely than those who did not report witnessing domestic violence to report alcohol use abuse; 1.4 odds ratio (95% CI 1.1 to 2.0). Witnessing domestic violence and cigarette use – The review identified one paper which reported on the association between witnessing domestic violence and cigarette use (Simantov et al., 2000 – reported for both males and females). The study found an association between witnessing domestic violence and cigarette use, with both female and male respondents who reported witnessing domestic violence being more likely to report cigarette use than those who did not report witnessing domestic violence; however this association was only found to be significant in females. The reported relative risk ratios ranged between 1.4 and 2.2. Simantov et al., 2000 (females): Female respondents who reported witnessing domestic violence were significantly more likely than those who did not report witnessing domestic violence to report cigarette use; relative risk ratio=2.2 (95% CI 1.6 to 3.2). Simantov et al., 2000 (males): Male respondents who reported witnessing domestic violence were more likely than those who did not report witnessing domestic violence to report cigarette use, but this association was not significant; Relative risk ratio=1.4 (95% CI 0.9 to 2.2). Witnessing domestic violence and ‘drug’ use/abuse – The review does not report on studies which measured the association between neglect and ‘drug’ use or abuse as it does for emotional abuse, physical	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Type of abuse - Prevalence estimates – Acierno et al., 2000/Kilpatrick et al., 2000 – Sexual assault 8%; physical assault 22%. Champion et al., 2004 – Sexual victimisation – male 2.8%; female 7.1%. Clark et al., 2004; 2005 – Supervisory neglect 6.5%. Fergusson et al., 1997; 1996 – Physical maltreatment – children reporting that they had 'regularly' been physically maltreated 7.6%; children reporting that had been physically maltreated 'too often' or 'too severely' 3.9%; sexual abuse 10.4%. Behnken et al., 2010 – Sexual abuse 11%. Bergen et al., 2004 – Sexual abuse – males aged 13, 14, or 15 = 1.6-2.0%; females aged 13, 14, or 15 = 5.4-6.7%. Chandy et al., 1997 – Sexual abuse 2.2%. Choquet et al., 1997 – Sexual abuse – rape 0.8%; attempted rape 2.1%; 'another sexual assault' 1.9%. Edgardh and Ormstad, 2000 – Sexual abuse – male 3.1%; female 11.2%. Erickson and Rapkin, 1991 – Unwanted sexual experiences 15%. Frederiksen et al., 2008 – Physical abuse – male 3.5%; female 2.7%. Garnefski and Arends,	abuse and sexual abuse. Witnessing domestic violence and use or abuse of other substances. The review does not report on studies which measured the association between witnessing domestic violence and use or abuse of other substances as it does for emotional abuse, physical abuse and sexual abuse (i.e. marijuana/hashish, cocaine/crack, methamphetamine, barbiturate, stimulants/uppers/speed, inhalants, hallucinogens, 'designer drugs', steroids, or medication/prescription). Emotional abuse and alcohol use/abuse – The review identified one paper which reported on the association between emotional abuse and alcohol use/abuse (Moran et al., 2004). The study found that there was a significant association between emotional abuse and alcohol use/abuse, with respondents who reported emotional abuse being significantly more likely than those who did not report emotional abuse to report alcohol use/abuse. Moran et al., 2004: Respondents who reported emotional abuse were significantly more likely than those who did not report emotional abuse to report alcohol use/abuse; 1.5 odds ratio (Reported by review authors as significant but 95% CI is not reported). Emotional abuse and cigarette use – The review identified one paper which reported on the association between emotional abuse and cigarette use (Moran et al., 2004). The study found that there was a significant association between emotional abuse and cigarette use, with respondents who reported emotional abuse being significantly more likely than those who did not report emotional abuse to report cigarette use.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	1998 – Sexual abuse 6%. Hamburger et al., 2008 – Sexual abuse 8.9%; physical abuse 22.3%, witnessing domestic violence 32.4%. Hernandez et al., 1992 – Sexual abuse (incest, extra familial) - male and female 10%; male 6.8%; physical abuse 9.1%. Hernandez et al., 1993 – Sexual abuse (incest, extra familial) - male and female 10%; male 6.8%; physical abuse 9.1%. Hibbard et al., 1988 – Physical abuse 10.3%; sexual abuse 4.1%. Hibbard et al., 1990 – Physical abuse 9.0%; sexual abuse 4.3%. Howard et al., 2005 – Sexual abuse – male 5.1%; female 10.2%. Lau et al., 2003 – Physical abuse – 'beaten for no reason/last 6 months' (p228) 10.9%; 'ever' beaten 10.4%. Logan et al., 2009 – Physical abuse – male 18%; female 19%. Luster and Small, 1997 – Sexual abuse – 'currently' 1%; 'before' 7%. Nagy et al., 1994 – Sexual abuse 12.6%. Nelson et al., 1994 – Sexual abuse 20.9%. Moran et al., 2004 – Emotional/verbal abuse 9.5%; physical abuse 10.6%; sexual abuse 5.5%. Pedersen and Skrondal, 1996 – Sexual abuse - male 1%; female 17%. Perkins and Jones,	Moran et al., 2004: Respondents who reported emotional abuse were significantly more likely than those who did not report emotional abuse to report cigarette use; 1.4 odds ratio (reported by review authors as significant but 95% CI is not reported). Emotional abuse and 'drug' use/abuse – The review identified one paper which reported on the association between emotional abuse and 'drug' use (Moran et al., 2004). The study found that the association was non-significant (statistical data not presented). Emotional abuse and use or abuse of other substances – The review did not identify any papers which reported on the associations between emotional abuse and use or abuse of marijuana or hashish; cocaine or crack; methamphetamines; barbiturates; stimulants (uppers, speed); inhalants; hallucinogens; 'designer drugs'; steroids; medication (prescription). Physical abuse and alcohol use/abuse – The review identified 14 papers which reported on the association between physical abuse and alcohol use/abuse (Fergusson et al., 1997; Frederikson et al., 2008; Hamburger et al., 2008; Hernandez et al., 1993; Hibbard et al., 1988; Hibbard et al., 1990; Kilpatrick et al., 2000; Lau et al., 2003; Moran et al., 2004; Perkins and Jones, 2004; Riggs et al., 1990; Shin et al., 2009; Southwick-Bensley et al., 1999; Yen et al., 2008). Both Frederikson et al., 2008 and Lau et al., 2003 used a number of different measures. Frederikson et al., 2008 measured associations in both females and males, whilst Southwick-Bensley et al., 1999 measured whether association varied by age group. All of	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	2004 – Physical abuse 20.1%. Riggs et al., 1990 – Sexual abuse 5.4%; physical abuse 5.2%. Shin et al., 2009 – Before grade 6 – physical abuse; sexual abuse; neglect (prevalence not reported). Simantov et al., 2000 – Family violence 25.9%. Southwick-Bensley et al., 1999 – Abuse 11%; sexual molestation 5.8%. Watts and Ellis, 1993 – Sexual molestation 8.7%. Yen et al., 2008 – Physical abuse 22.2%. • Looked after or adopted status - Not reported for any of the included studies. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported for any of the included studies. **Sample size:** The review included 35 articles covering 31 studies. The total sample size of the studies combined is not reported by the review authors. Due to concerns regarding overlaps between samples this has not been calculated by the NCCSC. Sample sizes for each study are given below: Acierno et al. 2000/Kilpatrick et al. 2000, n=4023 Behnken et al. 2010 n=6364 Bergen et al. 2004, n=2596 Champion et al. 2004,	the studies found at least 1 significant association between physical abuse and alcohol use/abuse, with respondents who reported physical abuse being significantly more likely than those who did not report physical abuse to report alcohol use/abuse. However, Frederikson et al., 2008 found that the significance of the association in females depended on the measure used; and Southwick-Bensley et al., 1999 found that significance of the association varied according to age group. Southwick-Bensley et al., 1999 also found a non-significant effect in the reverse direction; i.e. that participants who had experienced physical abuse were less likely to report alcohol use/abuse. The review authors suggest that this may be due to resilience or the result of protective factors such as foster care placement, extra-curricular activities, etc.). The reported odds/relative risk ratios of these studies ranged between 0.8 and 8.9. [19] Kilpatrick et al., 2000: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report alcohol use/abuse; odds ratio=3.9, 95% CI not reported; reported as significant by review authors. [24] Fergusson et al., 1997: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report alcohol use/abuse, statistical data not presented; reported as significant by review authors. [31] Frederikson et al., 2008 (results vary by measure; male): Male respondents who reported physical abuse were more likely than those who did not report physical abuse to report alcohol use/abuse; however this association was not significant; odds ratio=1.5 (95% CI 0.6 to 3.5). [31] Frederikson et al., 2008 (results vary by measure; male): Male respondents who reported physical abuse were	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	n=1883 Chandy et al. 1997, n=740 Choquet et al. 1997, n=183 Clark et al. 2004; 2005, n=170 Edgardh and Ormstad, 2000, n=1943 Erickson and Rapkin, 1991, n=1197 Fergusson et al. 1997; 1996, n=1025 Frederiksen et al. 2008 n=6009 Garnefski and Arends, 1998 n=1490 Hamburger et al. 2008 n=3559 Hernandez et al. 1992 n=3178 Hernandez et al. 1993 n=2973 Hibbard et al. 1988 n=712 Hibbard et al. 1990 n=3998 Howard et al. 2005 n=13601 Lau et al. 2003 n=489 Logan et al. 2009 n=1484 Luster and Small, 1997 n=36533 Moran et al. 2004 n=2164 Nagy et al. 1994. n=3018 Nelson et al. 1994 n=2332 Pedersen and Skrondal, 1996 n=597 Perkins and Jones, 2004 n=16313 Riggs et al. 1990 n=600 Shin et al. 2009 n=12478 Simantov et al. 2000 n=5513 Southwick-Bensley et al. 1999 n=4790 Watts and Ellis, 1993 n=670 Yen et al. 2008 n=1684. **Recognition indicators measured:** Substance abuse - (5 questions, prevalence rate). Substance use/abuse – Smoked cigarettes in last 30 days (prevalence rate); past year abuse/dependence on alco-	more likely than those who did not report physical abuse to report alcohol use/abuse; however this association was not significant; odds ratio=1.4 (95% CI 0.2 to 11.0). [31] Frederikson et al., 2008 (results vary by measure; female): Female respondents who reported physical abuse were more likely than those who did not report physical abuse to report alcohol use/abuse; however this association was not significant; odds ratio=1.1 (95% CI 0.6 to 1.8). [31] Frederikson et al., 2008 (results vary by measure; female): Female respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report alcohol use/abuse; odds ratio=8.9 (95% CI 2.5 to 32.1). [33] Hamburger et al., 2008: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report alcohol use/abuse; odds ratio=1.9 (95% CI 1.6 to 2.3). [35] Hernandez et al., 1993: Respondents who reported physical abuse were more likely than those who did not report physical abuse to report alcohol use/abuse; however this association was not significant, statistical data not presented; reported as non-significant by review authors. [36] Hibbard et al., 1988: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report alcohol use/abuse; relative risk ratio=1.8, 95% CI not reported; reported as significant by review authors. [37] Hibbard et al., 1990: Respondents who reported physical abuse were more likely than those who did not report physical abuse to report alcohol use/abuse; however this association was not significant, relative risk ratio=1.8 (95% CI 1.0 to 3.4). [39] Lau et al., 2003 (results vary with measure): Respondents who reported physical abuse were	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	hol/hard drugs/marijuana (prevalence rates). Kilpatrick et al., 2000 – Maltreatment - Sexual assault (4 questions, prevalence rate); physical assault (5 questions, prevalence rate). Substance use/abuse – Smoked cigarettes in last 30 days (prevalence rate); past year abuse/dependence on alcohol/hard drugs/marijuana (prevalence rate). Champion et al., 2004 – Maltreatment – Sexual victimisation (2 questions, self-reported, prevalence rate). Substance use/abuse – Alcohol – age of first drink; binge drinking in past 2 weeks (prevalence rate). Marijuana – Use in past 30 days. Clark et al., 2004/2005 – Maltreatment – Supervisory neglect (4 questions, prevalence rate). Substance use/abuse – Alcohol use disorder (prevalence rate). Self-report. Fergusson et al., 1997; 1996 – Maltreatment – Physical maltreatment (10 questions, prevalence rate); sexual abuse (prevalence rate). Substance use/abuse – Nicotine – dependence; alcohol –abuse/dependence; cannabis - abuse/dependence. Self-report. Behnken et al., 2010 – Maltreatment – Sexual abuse (1 question, self-report, prevalence rate). Substance use/abuse – Alcohol – 5 or more	significantly more likely than those who did not report physical abuse to report alcohol use/abuse; odds ratio=2.6, 95% CI not reported; reported as significant by review authors. [39] Lau et al., 2003 (results vary with measure): Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report alcohol use/abuse; odds ratio=2.9, 95% CI not reported; reported as significant by review authors. [44] Moran et al., 2004: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report alcohol use/abuse; odds ratio=2.1, 95% CI not reported; reported as significant by review authors. [46] Perkins and Jones, 2004: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report alcohol use/abuse, statistical data not presented; reported as significant by review authors. [47] Riggs et al., 1990: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report alcohol use/abuse; odds ratio=3.3 (95% CI 1.1 to 9.6). [48] Shin et al., 2009: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report alcohol use/abuse; odds ratio=1.3 (95% CI 1.0 to 1.8). [50] Southwick-Bensley et al., 1999 (results vary with age): Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report alcohol use/abuse; odds ratio=5.2 (95% CI 2.7 to 9.8). [50] Southwick-Bensley et al., 1999 (results vary with age): Respondents who reported physical abuse were more likely than those who did not report physical abuse to report alcohol use/abuse; however this	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	drinks in a row in past 30 days, prevalence rate. Bergen et al., 2004 – Maltreatment – Sexual abuse (1 question – yes/no – if yes, further 2 questions, prevalence rate). Substance use/abuse – Use in the last year of: - alcohol; tobacco; marijuana; acid or LSD; sniff glue, petrol or solvents; injected illegal drugs (heroin/speed); oral stimulants (speed, crack, ecstasy), magic mushrooms. Self-report. Chandy et al., 1997 – Maltreatment – Sexual abuse (1 question, prevalence rate). Substance use/abuse – Use of alcohol, tobacco and marijuana. Self-report. Choquet et al., 1997 – Maltreatment – Sexual abuse (1 question – if yes, further 3 questions, prevalence rate). Substance use/abuse – Consumption of alcohol, drugs, cigarettes. Smoking. Self-report. Edgardh and Ormstad, 2000 – Maltreatment – Sexual abuse (6 questions, prevalence rate, self-report). Substance use/abuse (self-report) – Alcohol – drunk before 15 years old; drugs – tried illicit drugs. Erickson and Rapkin, 1991 – Maltreatment – Unwanted sexual experiences (questions, prevalence rate). Substance use/abuse. Frederiksen et al., 2008 – Maltreatment (1 question – if yes, further 4, self-	association was not significant; odds ratio=1.7 (95% CI 0.8 to 3.4). [50] Southwick-Bensley et al., 1999 (results vary with age): Respondents who reported physical abuse were less likely than those who did not report physical abuse to report alcohol use/abuse; however this association was not significant; odds ratio=0.8 (95% CI 0.3 to 1.8). NB It appears that a mistake has been made in relation to the notation for this study. In the table a single X has been used but the legend uses an asterisk to denote that results vary with age. [52] Yen et al., 2008: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report alcohol use/abuse; odds ratio=3.4 (95% CI 2.4 to 4.7). Physical abuse and cigarette use – The review identified 8 papers which reported on the association between physical abuse and cigarette use (Acierno et al., 2000 – males and females; Fergusson et al., 1997; Frederikson et al., 2008 – males and females; Hibbard et al., 1988; Lau et al., 2003 – using two different measures; Moran et al., 2004; Perkins and Jones, 2004; and Riggs et al., 1990. All 8 studies found a significant association between physical abuse and cigarette use, with respondents who reported physical abuse being significantly more likely than those who did not report physical abuse to report cigarette use. Both Acierno et al., 2000 and Frederikson et al., 2008 found this to be the case in males and females; and Lau et al., 2003 found that this was the case using two different measures. The reported odds/relative risk ratios of these studies ranged between 1.8 and 6.1. [18] Acierno et al., 2000 (males): Male respondents who reported physical abuse were	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	report, prevalence rate). Substance use/abuse (self-report) – Frequency of use of alcohol and drugs. Garnefski and Arends, 1998 – Maltreatment – Sexual abuse (1 question, self-report, prevalence rate). Substance use/abuse (self-report) – Alcohol – use in last month; Marijuana – use in last year. Hamburger et al., 2008 – Maltreatment (self-report) – Sexual abuse (1 question, self-report, prevalence rate); physical abuse (1 question, prevalence rate); witnessing domestic violence (1 question, prevalence rate). Substance use/abuse – Ever drunk alcohol/age of initiation. Hernandez et al., 1992 – Maltreatment (self-report) – Sexual abuse – incest and extra familial (2 questions, prevalence rate); physical abuse (1 question, prevalence rate). Substance use/abuse (self-report) – Use of cigarettes and alcohol; frequency of intoxication and amount drunk; drinking problem; drug use (marijuana, inhalants, speed, cocaine, crack). Hernandez et al., 1993 – Maltreatment (self-report) – Sexual abuse – incest and extra familial (2 questions, prevalence rate); physical abuse (1 question, prevalence rate). Substance use/abuse (self-report) – Use of	significantly more likely than those who did not report physical abuse to report cigarette use; odds ratio=2.5 (95% CI not reported; reported as significant by review authors). [18] Acierno et al., 2000 (females): Female respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report cigarette use; odds ratio=4.1 (95% CI not reported; reported as significant by review authors). [24] Fergusson et al., 1997: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report cigarette use - statistical data not presented; reported as significant by review authors. [31] Frederikson et al., 2008 (males): Male respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report cigarette use; odds ratio=6.1 (95% CI 2.7 to 13.7). [31] Frederikson et al., 2008 (females): Female respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report cigarette use; odds ratio=2.8 (95% CI 1.6 to 4.8). [36] Hibbard et al., 1988 (relative risk ratio): Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report cigarette use; relative risk ratio=1.8 (95% CI not reported; reported as significant by review authors). [39] Lau et al., 2003 (results vary with measure): Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report cigarette use; odds ratio=3.2 (95% CI not reported; reported as significant by review authors, scale not reported). [39] Lau et al., 2003 (results vary with measure): Respondents who reported physical abuse were signifi-	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	cigarettes and alcohol; frequency of intoxication and amount drunk; drinking problem; drug use (marijuana, inhalants, speed, cocaine, crack). Hibbard et al., 1988 – Maltreatment (self-report) – Physical abuse (1 question, prevalence rate); Sexual abuse (1 question, prevalence rate). Substance use/abuse (self-report) – smoke cigarettes; ever used alcohol, marijuana or drugs. Hibbard et al., 1990 – Maltreatment (self-report) – Physical abuse (1 question, prevalence rate); Sexual abuse (1 question, prevalence rate). Substance use/abuse (self-report) – Use of alcohol and marijuana. Howard et al., 2005 – Maltreatment – Sexual abuse (1 question, prevalence rate). Substance use/abuse – Cigarettes – amount smoked in last 30 days; 5 or more drinks in last 30 days; use of cocaine or glue in last 30 days. Lau et al., 2003 – Maltreatment – Physical abuse – beaten for no reason/ever beaten (3 questions, prevalence rate). Substance use/abuse – Consumption of alcohol, smoking, or abuse of medication (prescription drugs) in last 6 months. Logan et al., 2009 – Maltreatment – Physical abuse (1 question, prevalence rate). Substance use/abuse – use of drugs	cantly more likely than those who did not report physical abuse to report cigarette use; odds ratio=4.4 (95% CI not reported; reported as significant by review authors, scale not reported). [44] Moran et al., 2004: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report cigarette use; odds ratio=2.3 (95% CI not reported; reported as significant by review authors). [46] Perkins and Jones, 2004: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report cigarette use - statistical data not presented; reported as significant by review authors. [47] Riggs et al., 1990: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report cigarette use; odds ratio=3.2 (95% CI 1.2 to 8.5). Physical abuse and drug use/abuse – The review identified 10 studies which reported on the association between physical abuse and 'drug' use/abuse (Kilpatrick et al., 2000 – using 2 different measures; Hernandez et al., 1993; Hibbard et al., 1988 – using two different measures; Hibbard et al., 1990; Lau et al., 2003 – using 2 different measures; Logan et al., 2009; Moran et al., 2004; Perkins and Jones; Riggs et al., 1990; Southwick-Bensley et al., 1999. All of these studies found a significant association between physical abuse and drug use/abuse, with participants who reported physical abuse being significantly more likely than those who did not report physical abuse to report drug use/abuse; with the exception of Riggs et al., 1990 which found that the association was not significant. Kilpatrick et al., 2000; Hibbard et al., 1988; and	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	(2-3 days per month) in last 12 months. Luster and Small, 1997 – Maltreatment – Sexual abuse (1 question, prevalence rate). Substance use/abuse – binge drinking. Nagy et al., 1994 – Maltreatment – Sexual abuse (3 questions, prevalence rate). Substance use/abuse – use of alcohol, use of drugs in past month. Nelson et al., 1994 – Maltreatment – Sexual abuse (2 questions, prevalence rate). Substance use/abuse – Smoke cigarettes; use of alcohol, marijuana in past 30 days; ever used cocaine. Moran et al., 2004 – Maltreatment – Emotional/verbal abuse (1 question, prevalence rate); physical abuse (1 question, prevalence rate); sexual abuse (1 question, prevalence rate). Substance use/abuse – chewing tobacco; frequency of smoking cigarettes; alcohol; drugs (marijuana, cocaine, barbiturates and heroin. Pedersen and Skrondal, 1996 – Maltreatment – Sexual abuse (1 question, prevalence rate). Substance use/abuse – alcohol consumption (alcohol related diagnosis by Rutgers Alcohol Problem Index – 23 questions). Perkins and Jones, 2004 – Maltreatment – Physical abuse (1 question, prevalence rate). Sub-	Lau et al., 2003 all found a significant association using more than one measure. The reported odds/relative risk ratios of these studies ranged between 1.8 and 20.4. [19] Kilpatrick et al., 2000 (results vary with measure): Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report drug use/abuse; odds ratio=4.8 (95% CI not reported; reported as significant by review authors). [19] Kilpatrick et al., 2000 (results vary with measure): Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report drug use/abuse; odds ratio=12.4 (95% CI not reported; reported as significant by review authors). [35] Hernandez et al., 1993: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report drug use/abuse - statistical data not presented; reported as significant by review authors. [36] Hibbard et al., 1988 (results vary with measure): Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report drug use/abuse; relative risk ratio=2.1 (95% CI not reported; reported as significant by review authors). [36] Hibbard et al., 1988 (results vary with measure): Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report drug use/abuse; relative risk ratio=2.1 (95% CI not reported; reported as significant by review authors). [37] Hibbard et al., 1990: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report drug use/abuse; relative risk ratio=2.2 (95% CI 1.1 to 4.4). [39] Lau et al., 2003 (results vary with measure): Respondents who reported	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	stance use/abuse – Use of tobacco, alcohol, drugs (e.g. amphetamines, crack/cocaine, and marijuana). Riggs et al., 1990 – Maltreatment – Sexual abuse (1 question – if yes, further 1, prevalence rate); physical abuse (1 question, prevalence rate). Substance use/abuse – Use of cigarettes, alcohol, drugs/marijuana. Shin et al., 2009 – Maltreatment (Self-report) – Physical abuse before grade 6 (1 question, prevalence rate); Sexual abuse before grade 6 (1 question, prevalence rate); neglect before grade 6 (1 question, prevalence rate). Substance use/abuse ((self-report) – Binge drinking (prevalence rate). Simantov et al., 2000 – Maltreatment (self-report) – Family violence (1 question, prevalence rate). Substance use/abuse (self-report) – Regular smoking of cigarettes or drinking of alcohol. Southwick-Bensley et al., 1999 – Maltreatment - Abuse (1 question, prevalence rate); molestation (1 question, prevalence rate). Substance use/abuse – Heavy or light alcohol or drug use (amphetamines, cocaine, hallucinogens, heroin, inhalants or other drugs, marijuana, tranquillisers, uppers/downers). Watts and Ellis,	physical abuse were significantly more likely than those who did not report physical abuse to report drug use/abuse; odds ratio=6.3 (95% CI not reported; reported as significant by review authors). [39] Lau et al., 2003 (results vary with measure): Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report drug use/abuse; odds ratio=20.4 (95% CI not reported; reported as significant by review authors). [40] Logan et al., 2009: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report drug use/abuse; prevalence ratio=1.8 (95% CI 1.3 to 2.3). [44] Moran et al., 2004: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report drug use/abuse; odds ratio=2.9 (95% CI not reported; reported as significant by review authors). [46] Perkins and Jones, 2004: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report drug use/abuse - statistical data not presented; reported as significant by review authors. [47] Riggs et al., 1990: Respondents who reported physical abuse were more likely than those who did not report physical abuse to report drug use/abuse, however this association was not significant; odds ratio=1.9 (95% CI 0.7 to 5.1). [50] Southwick-Bensley et al., 1999: Respondents who reported physical abuse were significantly more likely than those who did not report physical abuse to report drug use/abuse; odds ratio=2.6 (95% CI 1.7 to 4.9). Physical abuse and use or abuse of other substances – The review identified 4 papers which reported on	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	1993 – Maltreatment – Sexual molestation (questions, prevalence rate). Substance use/abuse (prevalence rate) – lifetime, past year, and last month use of: alcohol, cigarettes and drugs (barbiturates, cocaine, crack, designer drugs, hallucinogens, inhalants, marijuana, methamphetamine, narcotics, and steroids. Yen et al., 2008 – Maltreatment (self-report) – Physical abuse (1 question, prevalence rate). Substance use/abuse (self-report) – Alcohol – Ever drunk, ever been drunk in past month, ever '… experienced conflict with others or ever been blamed for drinking by your family' (p230). The authors note that many studies determined history of maltreatment simply by asking whether the participant had been abused which is acknowledged to limit the ability to '… capture the experience of maltreatment adequately' (p230). They also report that a number of studies failed to differentiate between substance use and abuse and that maltreatment was measured retrospectively.	the association between physical abuse and marijuana use/abuse (Fergusson et al., 1997; Hibbard et al., 1988; Hibbard et al., 1990; Kilpatrick et al., 2000). All 4 reported a significant association (statistical data not presented). The review identified 1 paper which reported on the association between physical abuse and medication (prescription) use/abuse (Lau et al., 2003). The study found a significant association (statistical data not presented, reported as significant by review authors). The review did not identify any papers which reported on the associations between physical abuse and use or abuse of cocaine or crack; methamphetamines; barbiturates; stimulants (uppers, speed); inhalants; hallucinogens; 'designer drugs'; steroids. Sexual abuse and alcohol use/abuse – The review identified 25 papers which reported on the association between sexual abuse and alcohol use/abuse. A number of studies used more than 1 measure and the results varied as a result; whilst others found that the association varied according to gender. The reported odds ratios or relative risk ratios ranged between 1.4 and 5.2 [19] Kilpatrick et al., 2000: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse; odds ratio=4.6 (95% CI not reported); reported as significant by the review authors. [20] Champion et al., 2004: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse; odds ratio=3.0 (95% CI 1.4 to 6.3); reported as significant by the review authors. [23] Fergusson et al., 1997 (results vary with measure): Respondents who reported sexual abuse were	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse; odds ratio=2.5 (95% CI1.3 to 4.7); reported as significant by the review authors. [23] Fergusson et al., 1997 (results vary with measure): Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse; odds ratio=3.3 (95% CI 1.7 to 6.6); reported as significant by the review authors. [23] Fergusson et al., 1997 (results vary with measure): Respondents who reported sexual abuse were more likely than those who did not report sexual abuse to also report alcohol use/abuse; but this association was not significant; odds ratio=1.9 (95% CI 0.8 to 4.7); reported as non-significant by the review authors. [25] Behnken et al., 2010: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse; odds ratio=2.1 (95% CI 1.8 to 2.5); reported as significant by the review authors. [26] Bergen et al., 2004 (results vary with measures): Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse (statistical data not presented; reported as significant by the review authors). [26] Bergen et al., 2004 (results vary with measures): Respondents who reported sexual abuse were more likely than those who did not report sexual abuse to also report alcohol use/abuse but this was not significant (statistical data not presented; reported as non-significant by the review authors). [27] Chandy et al., 1997: Respondents who reported sexual abuse were more likely than those who did not report sexual abuse to also report alcohol use/abuse	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		but this was not significant (statistical data not presented; reported as significant by the review authors). [28] Choquet et al., 1997 (females): Female respondents who reported sexual abuse were more likely than female respondents who did not report sexual abuse to also report alcohol use/abuse but this was not significant (statistical data not presented; reported as significant by the review authors). [28] Choquet et al., 1997 (males): Male respondents who reported sexual abuse were significantly more likely than male respondents who did not report sexual abuse to also report alcohol use/abuse (statistical data not presented; reported as significant by the review authors). [29] Edgardh and Ormstad, 2000: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse (statistical data not presented; reported as significant by the review authors). [30] Erickson and Rapkin, 1991: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse (statistical data not presented; reported as significant by the review authors). [32] Garnefski and Arends, 1998: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse (statistical data not presented; reported as significant by the review authors). [33] Hamburger et al., 2008: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse; odds ratio=1.9 (95% CI 1.4 to 2.5); reported as significant by the review authors. [34] Hernandez et al., 1992: Respondents who reported sexual abuse were significantly more likely than those who did not report	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		sexual abuse to also report alcohol use/abuse (statistical data not presented; reported as significant by the review authors). [35] Hernandez et al., 1993: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse (statistical data not presented; reported as significant by the review authors). [36] Hibbard et al., 1988: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse; relative risk ratio=2.3 (95% CI not reported); reported as significant by the review authors. [37] Hibbard et al., 1990: Respondents who reported sexual abuse were more likely than those who did not report sexual abuse to also report alcohol use/abuse but this was not significant (statistical data not presented; reported as significant by the review authors). [38] Howard et al., 2005 (results vary with measure): Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse; odds ratio=2.2 (95% CI 1.8 to 2.8); reported as significant by the review authors. [38] Howard et al., 2005 (results vary with measure): Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse; odds ratio=3.5 (95% CI 2.7 to 4.7); reported as significant by the review authors. [41] Luster and Small, 1997: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse (statistical data not presented; reported as significant by the review authors). [42] Nagy et al., 1994: Respondents who reported sexual abuse were significantly more	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		likely than those who did not report sexual abuse to also report alcohol use/abuse (statistical data not presented; reported as significant by the review authors). [43] Nelson et al., 1994: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse (statistical data not presented; reported as significant by the review authors). [44] Moran et al., 2004: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse; odds ratio=3.1 (95% CI not reported); reported as significant by the review authors. [45] Pederson and Skrondal, 1996: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse (statistical data not presented; reported as significant by the review authors). [47] Riggs et al., 1990: Respondents who reported sexual abuse were more likely than those who did not report sexual abuse to also report alcohol use/abuse but this was not significant; odds ratio=1.4 (95% CI 0.5 to 3.6); reported as non-significant by the review authors. [48] Shin et al., 2009: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse; odds ratio=2.3 (95% CI 1.2 to 4.4); reported as significant by the review authors. [50] Southwick-Bensley et al., 1999 (results vary with measure): Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse; odds ratio=3.8 (95% CI 1.2 to 12.4); reported as significant by the review authors. [50] Southwick-Bensley et al., 1999 (results vary with measure): Respondents who reported sexual abuse	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse; odds ratio=5.2 (95% CI 2.7 to 9.8); reported as significant by the review authors. [50] Southwick-Bensley et al., 1999 (results vary with measure): Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse; odds ratio=1.8 (95% CI 1.1 to 3.0); reported as significant by the review authors. [51] Watts and Ellis, 1993: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report alcohol use/abuse (statistical data not presented; reported as significant by the review authors). Sexual abuse and cigarette use – The review identified 11 papers which reported on the association between sexual abuse and cigarette use (Acierno et al., 2000 - males and females; Bergen et al., 2004; Chandy et al., 1997; Choquet et al., 1997; Hernandez et al., 1992; Hibbard et al., 1988; Howard et al., 2005 – results vary with measure; Nelson et al., 1994; Moran et al., 2004; Riggs et al., 1990; and Watts and Ellis, 1993. All studies are reported as finding a significant association, with the exception of Riggs et al., 1990 which found a non-significant association in the reverse direction (i.e. that participants who had experienced sexual abuse were less likely to report cigarette use. The review authors suggest that this may be due to resilience or the result of protective factors such as foster care placement, extra-curricular activities, etc.). The reported odds ratios or relative risk ratios ranged between 2.0 and 4.2. [18] Acierno et al., 2000 (males): Male respondents who reported sexual	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		abuse were significantly more likely than male respondents who did not report sexual abuse to also report cigarette use; odds ratio=2.4; 95% CI not reported; reported as significant by review authors. [18] Acierno et al., 2000 (females): Female respondents who reported sexual abuse were significantly more likely than female respondents who did not report sexual abuse to also report cigarette use; odds ratio=4.2; 95% CI not reported; reported as significant by review authors. [26] Bergen et al., 2004: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report cigarette use (statistical data not presented; reported as significant by review authors). [27] Chandy et al., 1997: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report cigarette use (statistical data not presented; reported as significant by review authors). [28] Choquet et al., 1997: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report cigarette use (statistical data not presented; reported as significant by review authors). [34] Hernadnez et al., 1992: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report cigarette use (statistical data not presented; reported as significant by review authors). [36] Hibbard et al., 1988: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report cigarette use; relative risk ratio=3.1; 95% CI not reported; reported as significant by review authors. [38] Howard et al., 2005 (results vary with measure): Respondents who reported sexual abuse were significantly more	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		likely than those who did not report sexual abuse to also report cigarette use; odds ratio=2.0 (95% CI 1.6 to 2.5); reported as significant by review authors. [38] Howard et al., 2005 (results vary with measure): Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report cigarette use; odds ratio=4.1 (95% CI 3.1 to 5.3); reported as significant by review authors. [43] Nelson et al., 1994: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report cigarette use (statistical data not presented; reported as significant by review authors). [44] Moran et al., 2004: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report cigarette use; odds ratio=3.0; 95% CI not reported; reported as significant by review authors. [47] Riggs et al., 1990 – Respondents who reported sexual abuse were less likely than those who did not report sexual abuse to also report cigarette use but this was non-significant; odds ratio=0.9 (95% CI 0.4 to 2.4); reported as non-significant by review authors. [51] Watts and Ellis, 1993: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report cigarette use (statistical data not presented; reported as significant by review authors). Sexual abuse and 'drug' use – 15 papers reported on the association between sexual abuse and 'drug' use/abuse (Bergen et al., 2004; Champion et al., 2004 – two samples; Choquet et al., 1997 – males and females; Edgardh and Ormstad, 2000 – males and females; Erickson and Rapkin, 1991; Hernandez	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		et al., 1993; Hibbard et al., 1988; Howard et al., 2005; Kilpatrick et al., 2000;Nagy et al., 1994; Moran et al., 2004; Riggs et al., 1990; Southwick-Bensley et al., 1999; Watts and Ellis, 1993). The reported odds ratios of the included studies ranged between 1.0 and 8.6. [19] Kilpatrick et al., 2000 (results vary with measure): Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report 'drug' use/abuse; odds ratio=3.8; 95% CI not reported, reported as significant by review authors. [19] Kilpatrick et al., 2000 (results vary with measure): Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report 'drug' use/abuse; odds ratio=8.6; 95% CI not reported, reported as significant by review authors. [20] Champion et al., 2004 (sample 2): Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report 'drug' use/abuse; odds ratio=2.8 (95% CI 1.3 to 5.7), reported as significant by review authors. [20] Champion et al., 2004 (sample 1): Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report 'drug' use/abuse; odds ratio=2.0 (95% CI 1.1 to 3.7), reported as significant by review authors. [26] Bergen et al., 2004: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report 'drug' use/abuse (statistical data not presented, reported as significant by review authors). [28] Choquet et al., 1997 (males): Male respondents who reported sexual abuse were significantly more likely than male respondents who did not report sexual abuse to also report 'drug' use/abuse (statistical data not presented,	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		reported as significant by review authors). [28] Choquet et al., 1997 (females): Female respondents who reported sexual abuse were more likely than female respondents who did not report sexual abuse to also report 'drug' use/abuse but this was not significant (statistical data not presented, reported as non-significant by review authors). [29] Edgardh and Ormstad, 2000 (males): Male respondents who reported sexual abuse were more likely than male respondents who did not report sexual abuse to also report 'drug' use/abuse but this was not significant (statistical data not presented, reported as non-significant by review authors). [29] Edgardh and Ormstad, 2000 (females): Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report 'drug' use/abuse (statistical data not presented, reported as significant by review authors). [30] Erickson and Rapkin, 1991: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report 'drug' use/abuse (statistical data not presented, reported as significant by review authors). [35] Hernandez et al., 1993: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report 'drug' use/abuse (statistical data not presented, reported as significant by review authors). [36] Hibbard et al., 1988 (results vary with measure): Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report 'drug' use/abuse; relative risk ratio=3.2; 95% CI not reported, reported as significant by review authors. [36] Hibbard et al., 1988: Respondents who reported sexual abuse were more likely than those who did not report sexual abuse to	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		also report 'drug' use/abuse but this was not significant (statistical data not presented, reported as non-significant by review authors. [37] Hibbard et al., 1990: Respondents who reported sexual abuse were more likely than those who did not report sexual abuse to also report 'drug' use/abuse but this was non-significant; relative risk ratio=1.7 (95% CI 0.8 to 3.6). [38] Howard et al., 2005 (results vary with measure): Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report 'drug' use/abuse; odds ratio=4.8 (95% CI 3.7 to 6.3), reported as significant by review authors. [42] Nagy et al., 1994: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report 'drug' use/abuse (statistical data not presented, reported as significant by review authors). [44] Moran et al., 2004: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report 'drug' use/abuse; odds ratio=3.9; 95% CI not reported, reported as significant by review authors. [47] Riggs et al., 1990: Respondents who reported sexual abuse were more likely than those who did not report sexual abuse to also report 'drug' use/abuse but this was non-significant; odds ratio=1.0 (95% CI 0.4 to 2.7), reported as significant by review authors. [50] Southwick-Bensley et al., 1999: Respondents who reported sexual abuse were significantly more likely than those who did not report sexual abuse to also report 'drug' use/abuse; odds ratio=3.3 (95% CI 2.1 to 5.3), reported as significant by review authors. [51] Watts and Ellis, 1993: Statistical data not presented, reported as both significant and non-significant suggesting that results may have varied by subsample or	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		measures used. The authors include a notation with this study although it is not clear what is meant or whether it relates to these finding, it simply states that '… specific types of drugs may not always be significant although overall results' (p231). Statistical data were not presented. Sexual abuse and marijuana or hashish use/abuse – The review identified 9 papers which reported on the association between sexual abuse and marijuana or hashish use/abuse. Eight of these found a significant association (Champion et al., 2004; Chandy et al., 1997; Garnefski and Arends, 1998; Hernandez et al., 1992; Kilpatrick et al., 2000; Nelson et al., 1994; Watts and Ellis, 1993); whilst the ninth found a non-significant association (Hibbard et al., 1988). NB. Statistical data not presented for any papers, significance reported by review authors. Sexual abuse and cocaine or crack use/abuse – The review identified 3 papers which reported on the association between sexual abuse and crack use/abuse. The association was found to be significant in Hernandez et al., 1992 and Nelson et al., 1994. Watts and Ellis, 1993 also found a significant association; however it appears that this study also found a non-significant association suggesting that results may have varied by subsample or measures used. The authors include a notation with this study although it is not clear what is meant or whether it relates to these finding, it simply states that '… specific types of drugs may not always be significant although overall results' (p231). Statistical data were not presented.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Sexual abuse and methamphetamine use/abuse – The review identified 1 paper which reported on the association between sexual abuse and methamphetamine use/abuse (Watts and Ellis, 1993). The association was found to be significant (statistical data not presented). Sexual abuse and barbiturate use/abuse – The review identified 1 paper which reported on the association between sexual abuse and barbiturate use/abuse (Watts and Ellis 1993). The association was found to be both significant and non-significant suggesting that results may have varied by subsample or measures used. The authors include a notation with this study although it is not clear what is meant or whether it relates to these finding, it simply states that "… specific 'types of drugs may not always be significant although overall results' (p231). Statistical data were not presented. Sexual abuse and stimulants (uppers, speed) use/abuse – The review identified 1 paper which reported on the association between sexual abuse and stimulant use/abuse (Hernandez et al., 1992). The association was found to be significant (statistical data not presented). Sexual abuse and inhalants use/abuse – The review identified 3 papers which reported on the association between sexual abuse and inhalant use/abuse (Bergen et al., 2004; Hernandez et al., 1992; Watts and Ellis, 1993). All 3 found a significant association (statistical data not presented).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Sexual abuse and hallucinogens use/abuse – The review identified 2 papers which reported on the association between sexual abuse and hallucinogens use/abuse (Bergen et al., 2004; Watts and Ellis, 1993). Both found a significant association (statistical data not presented). Sexual abuse and 'designer drugs' use/abuse – The review identified 1 paper which reported on the association between sexual abuse and 'designer drug' use/abuse (Watts and Ellis, 1993). The association was found to be both significant and non-significant suggesting that results may have varied by subsample or measures used. The authors include a notation with this study although it is not clear what is meant or whether it relates to these finding, it simply states that '… specific types of drugs may not always be significant although overall results' (p231). Statistical data were not presented. Sexual abuse and steroid use/abuse – The review identified 1 paper which reported on the association between sexual abuse and steroid use/abuse (Watts and Ellis, 1993). The association was found to be significant (statistical data not presented). Sexual abuse and medication (prescription) use/abuse – The review did not identify any papers which reported on the association between sexual abuse and steroid use/abuse. Age of initiation of substance use/abuse – The review reports that 6 papers found that a history of maltreatment was associated with a younger age of initiation	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		of substance use/abuse. Champion et al., 2004; Bergen et al., 2004; Hamburger et al., 2008; Pederson and Skrondal, 1996 all found an association between sexual abuse in childhood and earlier use of substances/substance abuse. Clark et al., 2004/2005 (using the same sample) and Fergusson et al., 1997 both found that childhood experience of physical abuse, neglect, and witnessing domestic violence were also linked to earlier use/abuse of alcohol. (NB Statistical data not presented, and it is not clear whether associations were significant.) Cumulative effects of experience of more than 1 form of maltreatment – The review reports that 4 papers (Hamburger et al., 2008; Luster and Small, 1997; Moran et al., 2004; Shin et al., 2009) found that the association between physical and sexual abuse combined and alcohol use/abuse was greater than the association between either physical abuse or sexual abuse and alcohol use/abuse individually (no statistical data presented, and it is not clear whether associations were significant). The review reports that 1 paper (Moran et al., 2004) found that the association between physical and sexual abuse combined and cigarette use was greater than the association between either physical abuse or sexual abuse and cigarette use individually (no statistical data presented, and it is not clear whether associations were significant). The review reports that 1 paper (Moran et al., 2004) found the association between physical and sexual abuse combined and 'drug' use/abuse was greater than the association between either physical abuse or sexual abuse and 'drug' use/abuse individually (no statistical data presented, and it is not clear whether associations were significant).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Age of initiation of substance use/abuse – The review identified 6 papers that reported that a history of maltreatment was associated with a younger age of initiation of substance use/abuse. Champion et al., 2004; Bergen et al., 2004; Hamburger et al., 2008; Pederson and Skrondal, 1996 all found an association between sexual abuse in childhood and earlier use of substances or substance abuse. Clark et al., 2004/2005 (using the same sample) and Fergusson et al., 1997 both found that childhood experience of physical abuse, neglect, and witnessing domestic violence were also linked to earlier use/abuse of alcohol. (NB. No statistical data provided, and it is not clear whether associations were significant.) **Narrative findings** Neglect and alcohol use/abuse – The review identified 3 papers reporting on 2 studies which tested the association between neglect and alcohol use/abuse (Clark et al., 2004; Clark et al., 2005; Shin et al., 2009). All 3 studies found an association between neglect and alcohol use/abuse, with respondents who reported neglect being more likely than those who did not report neglect to also report alcohol use abuse; however this association was only found to be significant in Clark et al., 2004 and Clark et al., 2005. The reported odds ratios of these studies ranged between 1.2 and 21.2. Neglect and cigarette use – The review did not identify any papers which reported on the association between neglect and cigarette use.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Neglect and 'drug' use/abuse – The review does not report on studies which measured the association between neglect and 'drug' use or abuse as it does for emotional abuse, physical abuse and sexual abuse. Neglect and use or abuse of other substances – The review does not report on studies which measured the association between neglect and use or abuse of other substances as it does for emotional abuse, physical abuse and sexual abuse (i.e. marijuana/hashish, cocaine/crack, methamphetamine, barbiturate, stimulants/uppers/speed, inhalants, hallucinogens, 'designer drugs', steroids, or medication/prescription). Witnessing domestic violence and alcohol use/abuse – The review identified 2 papers which reported on the association between witnessing domestic violence and alcohol use/abuse (Hamburger et al., 2008; and Simantov et al., 2000 – reported for both males and females). Both studies found an association between witnessing domestic violence and alcohol use/abuse, with respondents who reported domestic violence being more likely to report alcohol use/abuse than those who did not report witnessing domestic violence (for both females and males as reported in Simantov et al., 2000). Hamburger et al., 2008 found that the association was significant; whilst Simantov et al., 2000 found that the association was significant in females but not in males. The reported odds ratios of these studies ranged between 1.4 and 1.9. Witnessing domestic violence – Cigarette use – The review identified 1 paper which reported on the association between witnessing domestic violence and	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		cigarette use (Simantov et al., 2000 – reported for both males and females). The study found an association between witnessing domestic violence and cigarette use, with both female and male respondents who reported witnessing domestic violence being more likely to report cigarette use than those who did not report witnessing domestic violence; however this association was only found to be significant in females. The reported relative risk ratios ranged between 1.4 and 2.2. Witnessing domestic violence and 'drug' use/abuse – The review does not report on studies which measured the association between neglect and 'drug' use or abuse as it does for emotional abuse, physical abuse and sexual abuse. Witnessing domestic violence and use or abuse of other substances – The review does not report on studies which measured the association between witnessing domestic violence and use or abuse of other substances as it does for emotional abuse, physical abuse and sexual abuse (i.e. marijuana/hashish, cocaine/crack, methamphetamine, barbiturate, stimulants/uppers/speed, inhalants, hallucinogens, 'designer drugs', steroids, or medication/prescription). Emotional abuse and alcohol use/abuse – The review identified 1 paper which reported on the association between emotional abuse and alcohol use/abuse (Moran et al., 2004). The study found that there was a significant association between emotional abuse and alcohol use/abuse, with respondents who reported emotional abuse being significantly more likely than	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		those who did not report emotional abuse to report alcohol use/abuse. Emotional abuse and cigarette use – The review identified 1 paper which reported on the association between emotional abuse and cigarette use (Moran et al., 2004). The study found that there was a significant association between emotional abuse and cigarette use, with respondents who reported emotional abuse being significantly more likely than those who did not report emotional abuse to report cigarette use. Emotional abuse and 'drug' use/abuse – The review identified 1 paper which reported on the association between emotional abuse and 'drug' use (Moran et al., 2004). The study found that the association was non-significant (statistical data not presented). Emotional abuse and use or abuse of other substances – The review did not identify any papers which reported on the associations between emotional abuse and use or abuse of marijuana or hashish; cocaine or crack; methamphetamines; barbiturates; stimulants (uppers, speed); inhalants; hallucinogens; 'designer drugs'; steroids; medication (prescription). Physical abuse and alcohol use/abuse – The review identified 14 papers which reported on the association between physical abuse and alcohol use/abuse, (Fergusson et al., 1997; Frederikson et al., 2008; Hamburger et al., 2008; Hernandez et al., 1993; Hibbard et al., 1988; Hibbard et al., 1990; Kilpatrick et al., 2000; Lau et al., 2003; Moran et al., 2004; Perkins and Jones, 2004; Riggs et al., 1990; Shin et al., 2009;	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Southwick-Bensley et al., 1999; and Yen et al., 2008). Both Frederikson et al., 2008 and Lau et al., 2003 used a number of different measures. Frederikson et al., 2008 measured associations in both females and males, whilst Southwick-Bensley et al., 1999 measured whether association varied by age group. All of the studies found at least 1 significant association between physical abuse and alcohol use/abuse, with respondents who reported physical abuse being significantly more likely than those who did not report physical abuse to report alcohol use/abuse. However, Frederikson et al., 2008 found that the significance of the association in females depended on the measure used; and Southwick-Bensley et al., 1999 found that significance of the association varied according to age group. Southwick-Bensley et al., 1999 also found a non-significant effect in the reverse direction; i.e. that participants who had experienced physical abuse were less likely to report alcohol use/abuse. The review authors suggest that this may be due to resilience or the result of protective factors such as foster care placement, extra-curricular activities, etc.). The reported odds/relative risk ratios of these studies ranged between 0.8 and 8.9. Physical abuse and cigarette use – The review identified 8 papers which reported on the association between physical abuse and cigarette use (Acierno et al., 2000 – males and females; Fergusson et al., 1997; Frederikson et al., 2008 – males and females; Hibbard et al., 1988; Lau et al., 2003 – using two different measures; Moran et al., 2004; Perkins and Jones, 2004; and Riggs et al., 1990). All 8 studies found a significant association between physical	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		abuse and cigarette use, with respondents who reported physical abuse being significantly more likely than those who did not report physical abuse to report cigarette use. Both Acierno et al., 2000 and Frederikson et al., 2008 found this to be the case in males and females; and Lau et al., 2003 found that this was the case using two different measures. The reported odds/relative risk ratios of these studies ranged between 1.8 and 6.1. Physical abuse and drug use/abuse – The review identified 10 studies which reported on the association between physical abuse and 'drug' use/abuse (Hernandez et al., 1993; Hibbard et al., 1988 – using two different measures; Hibbard et al., 1990; Kilpatrick et al., 2000 – using two different measures; Lau et al., 2003 – using two different measures; Logan et al., 2009; Moran et al., 2004; Perkins and Jones; Riggs et al., 1990; Southwick-Bensley et al., 1999). All of these studies found a significant association between physical abuse and drug use/abuse, with participants who reported physical abuse being significantly more likely than those who did not report physical abuse to report drug use/abuse; with the exception of Riggs et al., 1990 which found that the association was not significant. Kilpatrick et al., 2000; Hibbard et al., 1988; and Lau et al., 2003 all found a significant association using more than one measure. The reported odds/relative risk ratios of these studies ranged between 1.8 and 20.4. Physical abuse and use or abuse of other substances – The review identified 4 papers which reported on the association between physical abuse and marijuana use/abuse (Fergusson et al., 1997; Hibbard et al.,	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		1988; Hibbard et al., 1990; Kilpatrick et al., 2000). All 4 reported a significant association (statistical data not presented). The review identified 1 paper which reported on the association between physical abuse and medication (prescription) use/abuse (Lau et al., 2003). The study found a significant association (statistical data not presented, reported as significant by review authors). The review did not identify any papers which reported on the associations between physical abuse and use or abuse of cocaine or crack; methamphetamines; barbiturates; stimulants (uppers, speed); inhalants; hallucinogens; 'designer drugs'; steroids. Sexual abuse and alcohol use/abuse – The review identified 25 papers which reported on the association between sexual abuse and alcohol use/abuse. A number of studies used more than one measure and the results varied as a result; whilst others found that the association varied according to gender. The reported odds ratios or relative risk ratios ranged between 1.4 and 5.2 Sexual abuse and cigarette use – The review identified 11 papers which reported on the association between sexual abuse and cigarette use (Acierno et al., 2000 - males and females; Bergen et al., 2004; Chandy et al., 1997; Choquet et al., 1997; Hernadnez et al., 1992; Hibbard et al., 1988; Howard et al., 2005 – results vary with measure; Nelson et al., 1994; Moran et al., 2004; Riggs et al., 1990; and Watts and Ellis, 1993). All studies are reported as finding a significant association, with the exception of Riggs et al., 1990 which found a non-significant association in the	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		reverse direction (i.e. that participants who had experienced sexual abuse were less likely to report cigarette use. The review authors suggest that this may be due to resilience or the result of protective factors such as foster care placement, extra-curricular activities, etc.). The reported odds ratios or relative risk ratios ranged between 2.0 and 4.2. Sexual abuse and 'drug' use – 15 papers reported on the association between sexual abuse and 'drug' use/abuse (Champion et al., 2004 – two samples; Choquet et al., 1997 – males and females; Kilpatrick et al., 2000; Bergen et al., 2004; Edgardh and Ormstad, 2000 – males and females; Erickson and Rapkin, 1991; Hernandez et al., 1993; Hibbard et al., 1988; Howard et al., 2005; Nagy et al., 1994; Moran et al., 2004; Riggs et al., 1990; Southwick-Bensley et al., 1999; and Watts and Ellis, 1993). The reported odds ratios of the included studies ranged between 1.0 and 8.6. Sexual abuse and marijuana or hashish use/abuse – The review identified 9 papers which reported on the association between sexual abuse and marijuana or hashish use/abuse. Eight of these found a significant association (Champion et al., 2004; Chandy et al., 1997; Garnefski and Arends, 1998; Hernandez et al., 1992; Kilpatrick et al., 2000; Nelson et al., 1994; Watts and Ellis, 1993); whilst the ninth found a non-significant association (Hibbard et al., 1988). NB. Statistical data not presented for any papers, significance reported by review authors.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Sexual abuse and cocaine or crack use/abuse – The review identified 3 papers which reported on the association between sexual abuse and crack use/abuse. The association was found to be significant in Hernandez et al., 1992; and Nelson et al., 1994. Watts and Ellis, 1993 also found a significant association; however it appears that this study also found a non-significant association suggesting that results may have varied by subsample or measures used. The authors include a notation with this study although it is not clear what is meant or whether it relates to these finding, it simply states that '… specific types of drugs may not always be significant although overall results' (p231). Statistical data were not presented. Sexual abuse and methamphetamine use/abuse – The review identified 1 paper which reported on the association between sexual abuse and methamphetamine use/abuse (Watts and Ellis, 1993). The association was found to be significant (statistical data not presented). Sexual abuse and barbiturate use/abuse – The review identified 1 paper which reported on the association between sexual abuse and barbiturate use/abuse (Watts and Ellis, 1993). The association was found to be both significant and non-significant suggesting that results may have varied by subsample or measures used. The authors include a notation with this study although it is not clear what is meant or whether it relates to these finding, it simply states that '… specific types of drugs may not always be significant although overall results' (p231). Statistical data were not presented.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Sexual abuse and stimulants (uppers, speed) use/abuse – The review identified 1 paper which reported on the association between sexual abuse and stimulant use/abuse (Hernandez et al., 1992). The association was found to be significant (statistical data not presented). Sexual abuse and inhalants use/abuse – The review identified 3 papers which reported on the association between sexual abuse and inhalant use/abuse (Bergen et al., 2004; Hernandez et al., 1992; Watts and Ellis, 1993). All 3 found a significant association (statistical data not presented). Sexual abuse and hallucinogens use/abuse – The review identified 2 papers which reported on the association between sexual abuse and hallucinogens use/abuse (Bergen et al., 2004; Watts and Ellis, 1993). Both found a significant association (statistical data not presented). Sexual abuse and 'designer drugs' use/abuse – The review identified 1 paper which reported on the association between sexual abuse and 'designer drug' use/abuse (Watts and Ellis, 1993). The association was found to be both significant and non-significant suggesting that results may have varied by subsample or measures used. The authors include a notation with this study although it is not clear what is meant or whether it relates to these finding, it simply states that '… specific types of drugs may not always be significant although overall results' (p231). Statistical data were not presented.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Sexual abuse and steroid use/abuse – The review identified 1 paper which reported on the association between sexual abuse and steroid use/abuse (Watts and Ellis, 1993). The association was found to be significant (statistical data not presented). Sexual abuse and medication (prescription) use/abuse – The review did not identify any papers which reported on the association between sexual abuse and steroid use/abuse. Age of initiation of substance use/abuse – The review reports that 6 papers found that a history of maltreatment was associated with a younger age of initiation of substance use/abuse. Champion et al., 2004; Bergen et al., 2004; Hamburger et al., 2008; Pederson and Skrondal, 1996 all found an association between sexual abuse in childhood and earlier use of substances or substance abuse. Clark et al., 2004/2005 (using the same sample) and Fergusson et al., 1997 both found that childhood experience of physical abuse, neglect, and witnessing domestic violence were also linked to earlier use/abuse of alcohol. (NB Statistical data not presented, and it is not clear whether associations were significant). Cumulative effects of experience of more than one form of maltreatment – The review reports that 4 papers (Hamburger et al., 2008; Luster and Small, 1997; Moran et al., 2004; Shin et al., 2009) found that the association between physical and sexual abuse combined and alcohol use/abuse was greater than the association between either physical abuse or sexual abuse and alcohol use/abuse individually (no statisti-	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		cal data presented, and it is not clear whether associations were significant). The review reports that 1 paper (Moran et al., 2004) found that the association between physical and sexual abuse combined and cigarette use was greater than the association between either physical abuse or sexual abuse and cigarette use individually (no statistical data presented, and it is not clear whether associations were significant). The review reports that 1 paper (Moran et al., 2004) found the association between physical and sexual abuse combined and 'drug' use/abuse was greater than the association between either physical abuse or sexual abuse and 'drug' use/abuse individually (no statistical data presented, and it is not clear whether associations were significant). Age of initiation of substance use/abuse – The review identified 6 papers that reported that a history of maltreatment was associated with a younger age of initiation of substance use/abuse. Bergen et al., 2004; Champion et al., 2004; Hamburger et al., 2008; Pederson and Skrondal, 1996 all found an association between sexual abuse in childhood and earlier use of substances or substance abuse. Clark et al., 2004/2005 (using the same sample) and Fergusson et al., 1997 both found that childhood experience of physical abuse, neglect, and witnessing domestic violence were also linked to earlier use/abuse of alcohol. (NB. No statistical data provided, and it is not clear whether associations were significant.)	

24. Wilson Steven R et al. (2010) Comparing physically abused, neglected, and nonmaltreated children during interactions with their parents: A meta-analysis of observational studies. Communication Monographs 77: 540–75

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To assess how abused and neglected children are distinguished from nonmaltreated children during interactions with their parents, on 3 behavioural clusters, communicating positivity (e.g., affection, approval), aversiveness (e.g., anger, resistance), and involvement (e.g., attention, interest). **Methodology:** Systematic review. **Country:** Range of countries. USA (26 studies), Spain (2 studies), Canada (2 studies). **Source of funding:** Not reported.	**Participants:** • Children and young people. Children with a documented history of being physically abused and/or neglected, compared with children with no history of maltreatment. • Caregivers and families - parents of children. **Sample characteristics:** • Age - children: mean age ranged from 1.2–11.5 years. parents: average age ranged from 20–34 years. • Sex - children: ranged from 12–91% boys parents: males and females. • Ethnicity – Not reported. • Religion/belief – Not reported. • Disability – Not reported. • Long term health condition – Not reported. • Sexual orientation – Not reported. • Socioeconomic position – Not reported. • Type of abuse - child abuse and neglect. Maltreatment type, 21 of the 36 comparisons (58%) involve abused versus nonmaltreated children, 7 (20%) involve	**Statistical data -** Do maltreated and nonmaltreated children display different levels of positivity, aversiveness or involvement during interactions with their parents when child behaviours are coded by independent observers? If so, do the behavioural clusters that most accurately distinguish neglected from nonmaltreated children differ from those that distinguish physically abused children? Do these questions depend on family or study characteristics? Observed child's behaviours in 1. Positivity: maltreated children vs non-maltreated children - mean weighted effect size (d)=minimum 0.42 (95% CI 0.30 to 0.54) to maximum 0.45 (95% CI 0.32 to 0.58). Chi-Sq test significant for homogeneity p=0.07 for minimum estimates of d to p=0.02 for maximum estimates of d, suggesting probability of other factors at work. (based on 29 effects sizes from 19 studies, no. of children 1545). 2. Aversiveness maltreated children vs. non-maltreated children - mean weighted effect size (d)=minimum 0.31 (95% CI 0.16 to 0.46) to maximum 0.29 (95% CI 0.12 to 0.46). Chi-Sq test significant for homogeneity p<0.001 for minimum estimates of d to p< 0.001 for maximum estimates of d, suggesting high probability of moderating factors. (based on 29 effect sizes from 24 studies, no. of children 1868). 3. Involvement maltreated children vs. non-maltreated children - mean weighted effect size (d) = minimum 0.51 (95% CI 0.25 to 0.77) to maximum 0.55 (95% CI	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** + Clear search strategy, inclusion and exclusion criteria, and good description of methods for meta-analysis, but no explicit details on how the quality of included studies was assessed. See internal validity.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	neglected versus nonmaltreated children, and 8 (22%) compare a mixed maltreatment group versus nonmaltreated children. • Looked after or adopted status – N/A. • Unaccompanied asylum seeking, refugee or trafficked children – N/A. **Sample size:** Systematic reviews: number of studies - 30 observational studies on how abused and neglected children are distinguished from non-maltreated children during interactions with their parents. Search dates: January 1994 to August, 2008 (studies published between 1978 and 2006). No. of children observed: 2,221. Average sample size of studies: 62 children. **Recognition indicators measured:** Interaction with parents - Children's behaviours during interactions with their parents in 3 categories: 1. aggressiveness: e.g., active physical aggression, verbal aggression, vocal negative, and noncompliance. 2. Positivity: e.g. verbal and nonverbal. 3. Involvement: initiation and social interaction. Interactions with their parents (children and their parents complete	0.29 to 0.81). Chi-Sq test significant for homogeneity p<0.001 for minimum estimates of d to p<0.001 for maximum estimates of d, suggesting high probability of moderating factors are present. (based on 22 effect sizes from 17 studies, no. of children 1136). 4. Moderators: Positivity, physical abused children vs non-maltreated children, mean effect size (d) = 0.44 (based on 13 effect sizes, no of children 820); Neglected children vs. non-maltreated children, mean effect size (d)=0.51 (based on 6 effect size, no of children 344); This suggests child's positive behaviours is equally useful in distinguishing either physically abused or neglected children from nonmaltreated children. 5. Moderators: Aversiveness, physical abused children vs. non-maltreated children, mean effect size (d)=0.29 (based on 18 effect sizes, no of children 1142); Neglected children vs. non-maltreated children, mean effect size (d)=0.30 (based on 5 effect sizes, no of children 308). This suggests that maltreatment type has little impact on child aversiveness during interactions with parents, which indicates that aversive behaviour is no more or less useful in distinguishing between physically abused and nonmaltreated children as opposed to neglected versus nonmaltreated children. 6. Moderators: Involvement, physical abused children vs. non-maltreated children, mean effect size (d)=0.39 (based on 12 effect sizes, no of children 577); Neglected children vs. non-maltreated children, mean effect size (d)=0.75 (based on 6 effect sizes, no	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	multiple tasks with different levels of structure) observed by independent observers. Level of tasks: highly structured, moderately structured, unstructured.	of children 374). This suggests that lack of child involvement during interactions with parents is a better diagnostic sign for distinguishing neglected as opposed to abused from nonmaltreated children. 7. Other moderators: Child and parent age, positivity: mean child age inversely associated with effect size in child positivity (p=0.08) Child<4.5 years age (infants, toddlers, pre-schoolers): comparing maltreated children vs. non-maltreated children, mean effect size (d)=0.57 (based on 13 effect sizes); Child =/>4.5 years age: comparing maltreated children vs. non-maltreated children, mean effect size (d)=0.25 (based on 11 effect sizes); A larger difference in positivity was observed in younger children when compared with older children, during their interaction with parents. Positivity: mean parent age inversely associated with effect size in child positivity (p=0.06) Parent<30 years age of maltreated children vs. non-maltreated children, mean effect size (d)=0.64 (based on 7 effect sizes); Parent=/>30 years age of maltreated children vs non-maltreated children, mean effect size (d)=0.35 (based on 6 effect sizes); A larger difference in positivity was observed in younger parents (with younger children) when compared with older parents with older children, during their parent-children interaction. Aversiveness: observation lengths was associated with effects for child aversiveness (p<0.02). Mean length of observation =/<15 mins (mean effect size (d) 0.05, based on 10 effect sizes); 16–59 mins (mean effect size (d) 0.40, based on 9 effect sizes);=/> 60 mins (mean effect size (d) 0.50, based on 10 effect sizes): comparing maltreated children vs. non-maltreated children. This suggests studies with longer observation periods obtain bigger effects. Observation	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		lengths and setting: observation length on average is longer in home than in lab studies. Home studies had a slightly larger effects (d=0.37) than lab studies (d=0.29). Nevertheless, observation length still is correlated with effect sizes for aversive behaviour even after the effects of setting are controlled (p<0.03). Observation length also is positively associated with parent age (p<0.05), and marginally associated with child age (p=0.06). Parents and children in studies with longer observation lengths tend to be older than those in studies with shorter observation lengths. However, observation length is a stronger predictor of effect size magnitude for aversive child behavior than is setting (home vs. lab), and the impact of observation length is not diminished after the effect of setting is controlled. Involvement One study (an outlier) obtained a very large effect size for child involvement in comparisons of both abused and neglected versus nonmaltreated children (d=>2.00). With this one study included, the mean weighted effect size (d) was 0.94 for home setting (based on 8 effect sizes, no. of children 376), and d=0.29 for lab setting. Removing this study from analysis, home setting (6 effect size, 293 children) still obtained an effect size of d=0.54 when compared with lab settings, in child involvement. Task structure is not associated with effect size when this outlier was removed. **Narrative findings** This systematic review and meta-analysis (n=30 observational studies, involving 2,221 children) compared abused, neglected, and nonmaltreated children's behaviour during interactions with their parents, in terms of communicating positivity (e.g., affection, approval), aversiveness (e.g., anger, resistance),	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		and involvement (e.g., attention, interest). Compared with non-maltreated children, a medium effect size was observed for both positive behaviours as well as for involvement, but between 'small' and 'medium' for aversive behaviors in maltreated children. The magnitude of difference between maltreated and nonmaltreated children's behaviour depends on multiple factors (moderators) such as type of maltreatment (abuse versus neglect), type of child behaviour, parent/child age, and observation length and setting. Child's positive behaviours is equally useful in distinguishing either physically abused or neglected children from nonmaltreated children. Maltreatment type has little impact on child aversiveness during interactions with parents, suggesting that aversive behaviour is no more or less useful in distinguishing between physically abused and nonmaltreated children as opposed to neglected versus nonmaltreated children. Lack of child involvement during interactions with parents is a better diagnostic sign for distinguishing neglected as opposed to abused from nonmaltreated children. A larger difference in positivity was observed in younger children when compared with older children, during their interaction with parents. A larger difference in positivity was also observed in younger parents (with younger children) when compared with older parents with older children, during their parent-children interaction. For aversiveness, studies with longer observation periods obtain bigger effects. Parents and children in studies with longer observation lengths tend to be older than those in studies with shorter observation lengths. However, observation length is a stronger predictor of effect size magnitude for aversive child behaviour than is setting (home versus lab), and the impact of observation length is not	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		diminished after the effect of setting is controlled. Home setting obtained a larger effect size when compared with lab settings, in child involvement. Task structure is not associated with effect size.	

Review question 5 – What tools support effective recognition of child abuse and neglect, and the taking of proportionate action?

Review question 5 – Critical appraisal tables

1. **Hershkowitz I, Fisher S, Lamb ME et al. (2007) Improving credibility assessment in child sexual abuse allegations: The role of the NICHD investigative interview protocol. Child Abuse and Neglect 31: 99–110**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: To study whether investigative interviews which follow the National Institute of Child Health and Human Development (NICHD) protocol, compared to those conducted in an unstructured way, enable interviewers to make more reliable and accurate judgements on the credibility of children's statements about their alleged experiences of sexual abuse. **Description of theoretical approach?** Partly. Some description of different type of information achieved using NICHD, although not a clear link to why this should make credibility judgement more accurate.	**Was the exposure to the intervention and comparison as intended?** Yes. Non-protocol interviews were conducted before protocol introduced. **Was contamination acceptably low?** Yes. **Did either group receive additional interventions or have services provided in a different manner?** Not reported. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Partly. All outcomes based on expert opinion. Plausibility of allegations (whether there was independent evidence that they	**Does the study's research question match the review question?** Partly. Study examines effectiveness of NICHD interview protocol in improving accuracy of detection of abuse. This tool aims to support children in disclosing abuse - therefore relates to confirmation rather than initial recognition. **Has the study dealt appropriately with any ethical concerns?** No. There is no mention of ethical approval, nor of obtaining consent from the young people or their parents/carers to use the interview transcripts for the purposes of this study (which was not the original purpose for which the interviews were conducted).	**Overall assessment of internal validity:** - **Overall assessment of external validity:** - Lack of ethical approval or informed consent is a significant flaw. **Overall validity rating:** - Little information provided regarding characteristics of interviewees - therefore unclear if interviews were 'typical' or 'atypical' cases. Limited information on statistical tests of relative accuracy of judgements for protocol versus non-protocol interviews.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
How was selection bias minimised? Matched groups. **Was the allocation method followed?** Yes. **Is blinding an issue in this study?** Blinding not possible. Assume not possible for raters to be blind to protocol use, as they were reading interview transcripts and presumably will be able to detect different questions/interviewing style. **Did participants reflect target group?** Partly. Two types of participants in the study: Youth investigators rating the interviews - sample comprised all current youth investigators, so was representative Children and young people with whom interviews were conducted - no information given about their characteristics. **Were all participants accounted for at study conclusion?** Yes.	were 'true' or 'false') was determined by a panel of experts using 'ground truth' scale (Horowitz et al. 1995), which in this study had inter-rater reliability of 91% (kappa=0.88), using corroborating evidence including physical and medical evidence, witness and suspect statements. Inter-rater reliability of judgements of credibility of children's testimony was generally good (Cronbach alpha=0.764 for non-protocol interviews; 0.874 for protocol interviews). However, inter-rater reliability was low for ratings of implausible allegations for non-protocol interviews (Cronbach alpha=0.338). **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Not reported. **Was follow-up time meaningful?** Not reported. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Not	**Were service users involved in the study?** No. Service users involved as participants, but not in design or interpretation of results. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children and young people suspected to have experienced sexual abuse. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to recognition, although potentially some overlap with assessment. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **Does the study have a UK perspective?** No. Israeli study.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	reported. Characteristics of interviewees for protocol versus non-protocol interviews not given. **Was intention to treat (ITT) analysis conducted?** Not reported. Not applicable. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. **Were the estimates of effect size given or calculable?** Not reported, but calculable. **Were the analytical methods appropriate?** Partly. Unclear whether comparison of accuracy judgements has been calculated using Friedman or McNemar test. No post-hoc comparisons conducted to detect where differences between the groups arose. **Was the precision of intervention effects given or calculable? Were they meaningful?** No. **Do conclusions match findings?** Yes.		

2. Hershkowitz I, Lamb ME, Katz C (2014) Allegation rates in forensic child abuse investigations: Comparing the revised and standard NICHD protocols. Psychology, Public Policy and Law 20, 336–44

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: This study aimed to assess the likelihood that victims in corroborated cases of intra-familial child maltreatment would make an allegation during interview using a revised version of the National Institute of Child Health and Human Development (NICHD) Investigative Interview Standard Protocol (RP), compared with the Standard Protocol (SP). The SP has been mandatory in Israel since 1996. The authors assert that the SP is effective for interviewing co-operative or motivated victims but is less attuned to the circumstances which can make it difficult for children who are reluctant to discuss what has happened to them. The RP was therefore designed to better support children reluctant to disclose abuse or make allegations. **Description of theoretical approach?** Yes. Theory that interviews with children who it is suspected have been abused, need more rapport building and better emotional support than current practice, particularly to support children who are reluctant to disclose or make an allegation of	**Was the exposure to the intervention and comparison as intended?** Yes. **Was contamination acceptably low?** Yes. All RP interviews conducted after SP interviews finished. **Did either group receive additional interventions or have services provided in a different manner?** No. Sample in SP and RP interviews was comparable - only difference was the type of protocol followed and the identity of the interviewer which was controlled for in the analysis. **Were outcomes relevant?** Yes. Study aimed to compare (validated) allegation rates after interviews conducted - this is what they looked at. **Were outcome measures reliable?** Partly. SP has been used since 1996. RP has not been validated. No specific detail given about how allegations were defined and recorded. This is concerning, given that the presence	**Does the study's research question match the review question?** Partly. Study looks at a tool to support children to make allegations rather than for professional to recognise signs of abuse. Study therefore relates to confirmation rather than initial recognition. **Has the study dealt appropriately with any ethical concerns?** Yes. Interview strategies in both protocols designed to avoid false allegations. Experienced child interviewers. Allegations corroborated by independent evidence. Interviews not conducted by authors (authors as supervisors). **Were service users involved in the study?** No. This would not have been appropriate due to their age and alleged victim status. Protocols were designed to support collection of accurate information and prior knowledge of them may have affected the outcome of the interviews. **Is there a clear focus on the guideline topic?** Yes. Clearly	**Overall assessment of internal validity:** - **Overall assessment of external validity:** - Lack of ethical approval or informed consent is a significant flaw. **Overall validity rating:** - Very little information given on how presence of an allegation was determined - this is significant, as this is the key outcome measure for the study. More detail could also be given on how, when and where the interviews took place; on whether the sample is powerful enough and the quasi-experimental nature of the study prevents all criteria being fulfilled.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
abuse. Not to disclose or make an allegation of abuse. **How was selection bias minimised?** Quasi-experimental SP and RP groups were nonrandomised but were considered similar in characteristics. Between-groups differences were also accounted for by use of hierarchical linear modelling. **Was the allocation method followed?** Not reported. There was no detail about how cases were chosen - analysis conducted on existing data set. **Is blinding an issue in this study?** Blinding not possible. **Did participants reflect target group?** Yes. In order to be included in the study, sample had to meet particular criteria. **Were all participants accounted for at study conclusion?** Yes. Nature of inclusion criteria made participants dropping out unlikely (analysis conducted after interviews had taken place).	or absence of allegation was a key outcome in the study. There also little detail about how allegations were validated - just a description of the kinds of independent evidence used in general. **Were all outcome measurements complete?** Yes. Children who did and didn't make allegations were identified. **Were all important outcomes assessed?** Partly. All allegations corroborated to be included in the study. No detail on how children felt during different kinds of interview in order so hard to say whether different protocols caused any harm to children being interviewed. **Were there similar follow-up times in exposure and comparison groups?** Not reported. No follow up necessary. **Was follow-up time meaningful?** Not reported. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. Characteristics of exposure and comparison group very similar in	supports recognition and assessment of child abuse processes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children who have experienced abuse (corroborated allegations). **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Yes, although there is no detail on where exactly the interviews took place - interviewers from Department of Child Investigation. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Could relate to recognition or assessment. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Study relates to recognition and assessment processes potentially as it is about supporting a child to make a valid allegation. **Does the study have a UK perspective?** No. Israeli study.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	age, gender, type of abuse, relationship to suspect and prior disclosure, although this was not tested statistically. Adjustment made via hierarchical linear modelling. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. No power calculations or effect sizes reported. **Were the estimates of effect size given or calculable?** Not reported. **Were the analytical methods appropriate?** Yes. Appropriate statistical testing, all potential influencing variables controlled for. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. p values given for allegation rates in comparison with different characteristics. **Do conclusions match findings?** Yes.		

3. Louwers EC, Korfage I, Affourtit MJ et al. (2012) Effects of systematic screening and detection of child abuse in emergency departments. Pediatrics 130: 457–64

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: Stated aim is to investigate whether introducing screening and training of nurses increases the detection rate of child abuse. However, not all intervention sites implemented training for nurses. **Description of theoretical approach?** No. **How was selection bias minimised?** Unmatched groups Before and after design. The study does not compare the characteristics of children in the before versus the after group. **Was the allocation method followed?** Partly. Not a 'pure' before and after design, in that: - Hospitals A and D screened for abuse using different checklists in the 'before' period - Hospital B also carried out some screening (<3%) unclear what checklists - Hospital C 'the existing checklist was adapted by using the Escape Form but was not completely replaced by it'.	**Was the exposure to the intervention and comparison as intended?** No. Intervention is described as being the 'Escape form' checklist. However 3 of the hospitals in the sample (hospitals E, F, G) are described as using different 'checklists with similar content' (p458). These are described as 'control hospitals' but their data appears to contribute to the overall analysis of impact of screening on detection rate. - One hospital continued to use another checklist alongside the Escape form in the post-implementation phase (hospital C). - Two hospitals implemented training for staff (A and C) but the others did not. Furthermore, in all hospitals, 100% of cases were not screened in the post-implementation period, meaning that some of the children in the 'intervention' group did not receive the intervention. Unclear if there is an intervention additional to checklist only. Authors note that 'If one of the warning signs was marked, the nurse informed the physician who had the responsibility to evaluate the increased risk for child abuse and take action if	**Does the study's research question match the review question?** Yes. Investigates effectiveness of systematic screening tool on detection of child abuse and neglect. **Has the study dealt appropriately with any ethical concerns?** Partly. Ethical approval received from Medical Committee at the University Medical Centre Rotterdam. However, no discussion in methodology about potential risks to children who were not screened for abuse, and how these might be mitigated. **Were service users involved in the study?** No. Service users involved as participants but not in design or interpretation. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children at risk of abuse and neglect (as indicated by presentation at emergency department).	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity rating:** - Although the study states it is examining the effectiveness of the 'Escape' screening checklist, in fact participating hospitals were using a variety of checklists. The role of the 'control' hospitals is unclear, and appear to have been included in the main analysis, rather than as a comparator. Validation of cases by an expert panel is a relatively weak outcome measure - it would have been a stronger design if the outcome of the child abuse team investigations had been followed up.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is blinding an issue in this study? Blinding not possible. **Did participants reflect target group?** Yes. **Were all participants accounted for at study conclusion?** Yes.	necessary' (p458). It is unclear how this was carried out. **Was contamination acceptably low?** No. See above. **Did either group receive additional interventions or have services provided in a different manner?** No. **Were outcomes relevant?** Partly. Study reports on detected cases of child abuse, which are then verified by a panel of experts. However, the design would have been more robust if they had followed up the detected cases to ascertain if the suspicions of abuse were substantiated in these cases. **Were outcome measures reliable?** Partly. Expert judgement rather than result of CPS investigation. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** No. It is unclear from the study what the rate of 'false positives' from the screening tool is.	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to recognition. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Incidence of abuse and neglect. **Does the study have a UK perspective?** No. Dutch study. Unclear to what extent practice context is the same. Emergency departments likely to be similar to UK A&E services. Referrals where abuse or neglect suspected are made to child abuse teams: 'multidisciplinary teams that deal with child abuse policy and assist hospital staff when child abuse is suspected' (p459). Unclear whether these are similar to UK children's social care services/Multi-Agency Safeguarding Hubs, or an internal team within the hospital.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Were there similar follow-up times in exposure and comparison groups? Not reported. **Was follow-up time meaningful?** Not reported. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Not reported. **Was intention to treat (ITT) analysis conducted?** No. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Partly. No power calculations reported, however the study has a large sample size (n=104,028). **Were the estimates of effect size given or calculable?** Yes. Odds ratios. **Were the analytical methods appropriate?** No. Study does not compare detection rates in the pre- versus post-implementation phases, but instead focuses on detection in checklist versus non-checklist. However, given that some hospitals were already using		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	checklists, and some were using checklists other than the Escape form, it makes it difficult to ascertain the effectiveness of this instrument. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. Confidence intervals provided. **Do conclusions match findings?** Partly. Study findings are reported in terms of effectiveness of screening generally, rather than the Escape tool specifically. This is consistent with how the analysis was conducted, but at odds with the described intention of the study which was to examine the effectiveness of the Escape screening tool.		

Review question 5 – Findings tables

1. **Hershkowitz I, Fisher S, Lamb ME et al. (2007) Improving credibility assessment in child sexual abuse allegations: The role of the NICHD investigative interview protocol. Child Abuse and Neglect 31: 99–110**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To study whether investigative interviews which follow the National Institute of Child health and Human Development (NICHD) protocol, compared to	**Participants:** Children and young people. Interview data from 24 children and young people used in the study.	**Statistical data**: **1. Inter-rater reliability** When rating non-protocol interviews, the variance between participants judgements was	**Overall assessment of internal validity** -

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
those conducted in an unstructured way, enable interviewers to make more reliable and accurate judgements on the credibility of children's statements about their alleged experiences of sexual abuse. **Methodology:** This study used a matched control group analysis. Interviews studied were conducted in Israel between 1994 and 2001. Half the interviews studied were from after the protocol became mandatory in Israel in 1998 and were individually matched to the other half, interviews conducted by the same professionals prior to this, which did not use the protocol. Cases were matched on children's age; type of allegation and strength of validating evidence before the authors sought out the transcripts of the interviews. Interviews were selected for study, from a larger group of interviews, because there was clear evidence that allegations made during the interviews were either plausible or implausible. Allegations were first assessed by three experts using a 'ground truth' scale (ratings included 'very likely' or 'likely', leading to a plausible assessment, or 'unlikely' or 'very unlikely', leading to an implausible assessment)	No information given regarding their characteristics. Professionals/practitioners - 42 Israeli youth investigators who had no prior knowledge of the cases selected for them to judge. **Sample characteristics:** (Characteristics of professionals) • Age - Average 34.5 years of age (SD= 10.1). • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Children in the interviews studied were alleging sexual abuse. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** Comparison numbers - 12 interviews conducted prior to 1998, non-protocol guided. Intervention numbers - 12 interviews conducted after	wider and there was lower inter-rater reliability (Cronbach's alpha=.764) than when rating protocol interviews (Cronbach's alpha=.874). This was particularly true when rating cases with implausible allegations (Cronbach's alpha=.642 for protocol and Cronbach's alpha=.338 for non-protocol interviews) but was less evident when rating cases with plausible allegations (Cronbach's alpha=.811 for protocol and Cronbach's alpha=.890 for non-protocol interviews). **2. Use of 'no judgement possible' (NJP) option** 16.7% of judgements of protocol interviews involved the NJP category (including 4.8% of plausible protocol investigations) compared with 52.4% of the non-protocol interviews (57.1% of all plausible cases and 47.6% of implausible cases). A McNemar non-parametric test for related samples confirmed this finding, that NJP option was used more frequently with non-protocol interviews than protocol (p<.019). The effect was evident and nearly significant (p<.06) when only plausible cases were analysed but not when only implausible cases were analysed. **3. Accuracy of judgements** 59.5% of the ratings of protocol interviews (95.2% of plausible cases, 23.8% implausible cases) were accurate, compared with	**Overall assessment of external validity** - Lack of ethical approval or informed consent is a significant flaw. **Overall validity score** - Little information provided regarding characteristics of interviewees - therefore unclear if interviews were 'typical' or 'atypical' cases. Limited information on statistical tests of relative accuracy of judgements for protocol versus non-protocol interviews.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
(p103). The assessment was based on evidence which included: reports by disinterested witnesses who observed (or failed to observe) some or all alleged events; physical evidence that events could or could not have taken place (e.g. photographs or video); medical evidence (indicating existence or nonexistence of expected injury) and/or suspect statements which gave details of alleged events which matched the details given by the alleged victim. Twenty-four interviews with children alleging sexual abuse were selected on this basis - with independent evidence indicating they were plausible (n=12) or implausible (n=12). Half (n=6) of interviews in each category were protocol-guided and half were not. Forty-two experienced Israeli youth investigators were then asked to rate the credibility of allegations made by the children in the interviews, with 7 participants independently judging each transcript. Participants did not have access to the 'ground truth' evidence and had to make a judgement based only on the interviews. Each participant was asked to rate 4 interviews (1 protocol-guided plausible, 1 proto-	1998 when protocol became mandatory. Sample size -n=24 (12 protocol guided, 12 non-protocol guided). **Intervention**: The tool used was the NIHCD protocol. It is a framework for forensic interviews, designed to enable interviewers to retrieve complete and accurate accounts of abuse from young alleged victims and witnesses. It involves creating a supportive interviewing environment (the presubstantive rapport building phase), using language appropriate to the child's developmental level while avoiding interruptions, giving children ground rules so they know what is expected of them and what they can do during the interview, training them to use their episodic memory to describe events and using words and gestures which tap into children's free recall memory. Interviewers are trained to maximise their use of open-ended questions and probes, to use focused questions only when open-ended ones have been exhausted, to only use option-posing questions at the end of the interview, to obtain essential information and to avoid suggestive questioning. Delivered by - Professionals who originally conducted the interviews are not described in any real detail.	29.6% of judgements of non-protocol cases (38.1% of judgements about plausible cases, 11.9% of implausible). Both plausible and implausible allegations were significantly more likely to be judged accurately ($p<.0001$ in both cases, test statistics not reported) in interviews where the protocol was used, compared to those where it was not. Accuracy effect size calculated by the review team: Accurate plausible judgements: Large effect size in favour of protocol interviews (d=1.92) Accurate implausible judgements: Small to medium effect size in favour of protocol interviews (d=0.46).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
col-guided implausible, 1 non-protocol plausible, 1 non-protocol implausible). They used a 4 point scale - 'very likely' or 'likely', leading to a plausible assessment, or 'unlikely' or 'very unlikely', with a fifth No Judgement Possible (NJP) option. They were also asked to indicate on a 5-point scale, from 'very unconfident' to 'very confident' how much confidence they had in their judgements. 'Inter-rater reliability was assessed using Cronbach's Alpha following a matrix transposition of the data' (p103). Each participant's rating was compared with the plausibility decision made by the experts prior to the selection of the interview. The judgement was then analysed with the respect to the plausibility of the allegation in question and whether it was protocol-guided or not. This was analysed 'using Friedman's or McNemar's (for dichotomous variables) non-parametric tests for related samples' (p104). A comparison of participant's confidence in their judgements (across the use of protocol and plausibility of statement) was conducted using a within-subjects analysis of variance (ANOVA). 'Pearson correlation coefficients	Delivered to - Children alleging sexual abuse. Duration, frequency, intensity - 12 interviews which occurred between 1994 and 1998 were analysed over a non-described timeframe in 2003.These did not use the protocol. 12 further interviews which occurred between 1998 and 2001 were also analysed. Location/place of delivery - Location of interviews and analysis was not described. **Describe comparison intervention –** Non-protocol interviews were conducted according to the preferences of the interviewer, with no specific structure. **Outcomes measured:** Incidence of abuse and neglect. Judgements were based on plausibility of allegations first by 3 experts based on documentary evidence, before being rated by study participants solely on the contents of an interview. Participants rating solely on interview judged how likely it was that the alleged incidents really happened using a 4-point scale: 'very likely', 'quite likely' 'quite unlikely', 'very unlikely'.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
were used to assess the association between accuracy and confidence' (p104). **Country:** Not UK. Israel. **Source of funding**: Government - Israeli Department of Youth Corrections and Investigation.			

2. Hershkowitz I, Lamb ME, Katz C (2014) Allegation rates in forensic child abuse investigations: Comparing the revised and standard NICHD protocols. Psychology, Public Policy and Law 20: 336–44

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: This study aimed to assess the likelihood that victims in corroborated cases of intrafamilial child maltreatment would make an allegation during interview using a revised version of the National Institute of Child Health and Human Development (NICHD) Investigative Interview Standard Protocol (RP), compared with the Standard Protocol (SP). The SP has been mandatory in Israel since 1996. The authors assert that the SP is effective for interviewing co-operative or motivated victims but is less attuned to the circumstances which can make it difficult for children who are reluctant to discuss what has happened to them. The RP was therefore designed to better support children reluctant to	**Participants**: Children and young people. Children referred for investigation following suspicions of physical or sexual abuse by family members. All cases were corroborated by one or more forms of independent evidence, for example admission by a suspect or medical evidence. **Sample characteristics:** • Age -The sample was made up of children aged 4–13 years - the mean age was 8.08, with a standard deviation (SD) of 2.57. SP group: mean age=8.02 years, SD= 2.55 RP group: the mean age=8.11 years, SD=2.58. • Sex -The sample was made up of 232 boys and 194 girls. SP group: 90 boys and 75 girls RP group: 142 boys and 119 girls	**Statistical data:** **Association between protocol type and allegation rates** 1. Fisher's exact test Allegation rates were significantly higher (calculated using Fisher's exact test) when using RP (59.8%) compared to SP (50.3%) (p=0.035). 2. Spearman correlations Likelihood of making an allegation was significantly correlated with the type of protocol used in the interview (r=0.093, p<0.05). (Note: these are the r and p values reported in the study – r value seems somewhat low.) 3. Hierarchical linear modelling The logistical hierarchical linear model that was generated found that the protocol version significantly predicted	**Overall assessment of internal validity** - **Overall assessment of external validity** - Lack of ethical approval or informed consent is a significant flaw. **Overall validity score** - Very little information given on how presence of an allegation was determined - this is significant, as this is the key outcome measure

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
disclose abuse or make allegations. **Methodology**: The study used a non-random pre-post intervention design. All RP interviews were only conducted after SP interviews had been completed. Fisher's exact tests were used to compute the rate of allegations in whole sample; how the rate of allegations varied according to the interviewers' identity and how it varied according to which protocol version was used. Spearman's correlation tests were conducted for all variables: age, gender, relationship to suspect, previous reporting, abuse type, protocol version and whether or not an allegation was made. In order to ascertain whether the type of protocol could predict whether a child would make an allegation or not (controlling for other possible influences) logistic Hierarchical Linear Modelling (HLM) was also used. As age can influence whether a child does or doesn't make an allegation, it was treated as a categorical variable in the Fisher's tests and a continuous variable in the HLM. **Country**: Not UK. Israel. **Source of funding**: Voluntary/charity -	• Ethnicity - Not reported. Religion/belief -Not reported. • Disability - Not reported. Long term health condition - Not reported. • Sexual orientation - Not reported. Socioeconomic position - Not reported. Type of abuse - The children were referred for investigation following suspicions of physical (n=408; SP: 159, RP: 239) or sexual abuse (n=18, SP: 6, RP: 12) by family members, either biologically related (n=375; SP: 144, RP: 231) or stepparents and stepsiblings (n=51; SP: 21, RP: 30). For all children included in the study these suspicions were corroborated by independent external evidence. Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - No reference made but implied all children had Israeli nationality and lived with family. **Sample size**: Comparison numbers - 165 children were included who were interviewed using the SP. Intervention numbers - 261 children were included who were interviewed using the RP.	whether an allegation would be made, once the effects of other factors including interviewer's identity, were controlled for. Using the RP significantly increased the possibility that children would make an allegation (coefficient=0.450, SE =0.214, t=2.104, exponent coefficient=1.568, p <.036). The model correctly predicted 64.1% outcomes - 50.3% of those without allegations and 74.9% of those with allegations. There was a very small subsample of children (n=18) who alleged sexual abuse. The model was therefore tested again on the sample of those children who alleged physical abuse (n=408). Protocol version again significantly predicted allegation once other factors except type of abuse were controlled for, with p<0.046. Again, using the RP significantly increased the possibility that children would report allegations (Exponential coefficient=1.547, F(5,403)=7.077, p<0.001). The model correctly classified 64% - 49.7% of cases of physical abuse without allegations and 74.9% of those with allegations.	for the study. More detail could also be given on how, when and where the interviews took place; on whether the sample is powerful enough and the quasi-experimental nature of the study pre-vents all criteria being fulfilled.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Grant from the Nuffield Foundation with support from the Israeli Department of Child Investigation whose interviewers and supervisor worked on the project.	Sample size - 426 children (SP:165, RP:261) were included in the sample, out of a larger group, where their suspected intrafamilial abuse was 'corroborated by independent external evidence including suspects' admissions, disinterested eyewitness testimony, medical evidence (including observable physical injuries) and material evidence' (p339). Total of 496 boys and 502 girls were excluded from the study because there was no corroborating evidence recorded on file. **Intervention category**: Other. **Intervention:** The intervention was a revised version of a standard protocol, SP, for interviewing children who are suspected victims of maltreatment that has been used in Israel since 1996. Interviews using the revised protocol, the RP, were conducted after all interviews using the SP has been completed. Delivered by - 7 experienced child interviewers from all regions across Israel, who work for the Israeli Department of Child Investigation. Delivered to - In total, 1424 children were interviewed, but only 426 children were included in the study because their suspected victimisation was corroborated. Duration, frequency, intensity- There were in total 613 interviews conducted		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	using the SP which occurred over 8 months. Interviewers were then introduced to the RP in a training session of 2 days length before conducting 811 interviews using the RP over another 8 month period. All interviews, pre- and post-RP training were considered for inclusion. Interviewers received supervision from two of the authors on a monthly basis. Supervision during SP interviews focused on cognitive rather than socioemotional factors, while supervision of RP interviews focused only on the socioemotional factors, in order to ensure supervision sessions were of a similar length and reflected the style of interview. The results are drawn from the interviews with 426 children whose suspected victimisation was corroborated. Key components and objectives of intervention- The SP, the standard NICHD protocol for forensic child abuse investigations, was designed to improve the informativeness of children's accounts of abuse by looking at the cognitive, linguistic and social factors which can affect that. It gives non-suggestive assistance to children to generate and organise their experiences of abuse. It involves the following phases: 1. In the introductory phase, interviewers introduce themselves, clarify the		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	need for the children to describe events they have actually experienced and in detail, and explain that they can and should point out when they don't know, understand or remember something as well as correct the interviewers if needed. 2. In the rapport building phase, made of two sections, the first structured open-ended section is designed to encourage children to provide information that is meaningful to them. The second involves prompting the children to describe one recently experienced event in detail, so that they use their episodic memory, while enhancing further the rapport between interviewer and child. 3. The transitional phase involves the interviewer using open-ended prompts to identify the target events to be investigated. If the child has still to make an allegation of abuse, interviewers can narrow the prompts and make reference to other knowledge, such as previous disclosures or physical marks. 4. When an allegation has been made the substantive phase of investigation begins. The authors of this study suggest however, as a result of their research into interviews using the SP, that its structure does not support reluctant children and can involve interviewers putting pressure on reluctant children to answer questions, changing		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	topic of conversation to sensitive matters before the child is ready to do so and using 'intrusive' rather than open-ended prompts (p337). Their changes and additions, to create the RP, included: 1. Moving the rapport building phase prior to the section on what the children can and are expected to do during the interview, and giving additional guidance to interviewers on building and maintaining rapport. 2. During this phase, interviewers are encouraged to address the children by name; welcome them and express care and interest in the children's feelings and experiences; to echo acknowledge and explore children's feelings as well as asking them to discuss and provide more information about events or experiences meaningful to them. 3. Interviewers are advised to encourage children verbally and nonverbally to describe experiences in both the first three presubstantive phases and the substantive phase whilst also thanking, appreciating and providing positive reinforcement to the children about their efforts rather than the content. 4. Interviewers were advised to express empathy regarding the children's		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	feelings about the interview experience rather than past experiences. 5. Interviewers were encouraged to use other supportive interventions such as legitimising expression; generalising any difficulties the child might be having during the interview; offering help to make it easier, and offering reassurance or optimism that they can overcome any difficulty where necessary or possible. 6. Where the child does report abuse but admits reluctance about discussing it further, interviewers are advised to contain or encourage the children as well as emphasising that the experience is not the child's fault. A prompt by prompt analysis of both protocol versions conducted by the authors found that there was no difference in questioning style - important in the context of risks of false allegations if interviews are conducted suggestively. RP and SP both contained equally few recognition prompts (such as option posing or suggestive prompts) with recall strategies in both protocols being dominantly open-ended. Content/session titles- See description above. Location/place of delivery - Not reported.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Describe comparison intervention: See description above. **Outcomes measured:** Incidence of abuse and neglect. Rate of allegations corroborated by external evidence.		

3. **Louwers EC, Korfage I, Affourtit MJ et al. (2012) Effects of systematic screening and detection of child abuse in emergency departments. Pediatrics 130: 457–64**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: Stated aim is to investigate whether introducing screening and training of nurses increases the detection rate of child abuse. However, not all intervention sites implemented training for nurses. **Methodology:** Other: Authors describe as 'prospective intervention cohort study' (p458). However, individual participants are not followed up at successive time points. Rather, the study has elements of a quasi-experimental study - one group of patients in the pre-implementation phase receives one type of service, and a second group in the post-implementation phase receives a second type of service. **Country:** Not UK - The Netherlands.	**Participants:** Children and young people. Children and young people presenting at emergency departments. Seven emergency departments out of 22 in the Netherlands participated in the study. **Sample characteristics:** • Age – Whole sample: 0 to 4: 40% 5–8: 17% 9–12: 17% 13–18: 56% Figures for intervention and control groups not available. • Sex -Whole sample: Male: 56% Female: 44% Figures for intervention and control groups not available. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition -Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** Comparison numbers - Children not screened: 66,624.	**Statistical data:** As noted in the critical appraisal, the analysis does not compare the intervention with any comparison groups but rates of detection amongst screened compared to non-screened children - presumably this includes all types of screening. Overall detection rate during the 23-month period was 0.2%. Detection rate was significantly higher for screened children compared to those not screened (0.5% vs 0.1%, p<0.001). The pooled odds ratio for detection amongst screened compared to non-screened children across the seven hospitals was OR=4.88 (95% CI 3.58 to 6.68).	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity rating:** - Although the study states it is examining the effectiveness of the 'Escape' screening checklist, in fact participating hospitals were using a variety of checklists. The role of the 'control' hospitals is unclear, and appear to have been included in the main analysis, rather than as a comparator. Validation of cases by an expert panel is a relatively weak outcome measure - it would have been a stronger design if the outcome of the

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Intervention numbers - Children screened: 37,404. Sample size - 104,208. **Intervention**: 'Escape form' checklist developed on basis of systematic review (Louwers et al. 2010). Checklist with 6 questions on warning signs for child abuse . Authors note that 'If one of the warning signs was marked, the nurse informed the physician who had the responsibility to evaluate the increased risk for child abuse and take action if necessary' (p458). It is unclear how this was carried out. Delivered by - Emergency department nurses. **Describe comparison intervention** Three hospitals are described as 'control hospitals' and implemented 'checklists with similar content' (p458). It is unclear what these checklists were. **Outcomes measured:** Incidence of abuse and neglect. Positive screen results lead to suspected cases of child abuse. The notes of these cases were then evaluated by a panel of 4 professionals (1 forensic paediatrician, 2 social paediatricians and a physician) to determine if the case was a 'potential case' or 'no case'. All 4 professionals agreed in 50% of the scored cases, overall		child abuse team investigations had been followed up.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	agreement rate for 3 or more professionals was 70.6%. This has been included as a valid outcome as it involves an independent verification of the outcome of the tool. However, following up the outcomes of any child protection investigations would have provided a more robust design.		

Review question 6 – What aspects of professional practice support and hinder recognition of child abuse and neglect, and the taking of proportionate action?

Review question 6 – Critical appraisal tables

1. Allnock D, Miller P (2013) No one noticed, no one heard: a study of disclosures of childhood abuse. London: NSPCC

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Qualitative approach was appropriate given the sensitive nature of the topic. The research team carried out a 2-hour in-depth interview with participants and were mindful of participant fatigue during data collection, so avoided duplication of data collected through maltreatment study survey. **Is the study clear in what it seeks to do?** Clear. Clear research question relating to young adult's views on disclosure which	**Is the context clearly described?** Clear. Data is collected on the 60 participants and characteristics are provided in Chapter 2 of the report (pp13–16). Where findings are presented, the authors clearly state the sex of the young adult and the abuse experienced which contextualises the data. **Was the sampling carried out in an appropriate way?** Somewhat appropriate. The authors state that 'there were two different recruitment methods employed in the study which may have introduced	**Does the study's research question match the review question?** Yes. The study is in relation to young people's views and experiences of seeking help and what factors support and hinder recognition. **Has the study dealt appropriately with any ethical concerns?** Yes. The study includes a thorough, comprehensive ethics section (p58), which state that the researchers who conduct the interviews with the young people are knowledgeable about child safety	**Overall assessment of internal validity:** ++ All of the criteria have been fulfilled to a high standard. The author has met all ethical concerns and sought solutions for young adults who may be affected by their participation in the study. There is a large sample group (n=60) and their characteristics are clearly described in the report. Findings prompt implications for practice so as to tailor better support and intervention for childhood abuse. The limitations are defined

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
add depth to childhood experiences of abuse. **How defensible/rigorous is the research design/methodology?** Defensible. The authors are explicit in the recruitment of participants and interview sample explaining that the majority were recruited via the NSPCC. Participants were given a £30 high street vouchers as a thank-you for taking part. The analysis included recording to be digitalised via NVivo, given permission from participants which was fully transcribed and anonymised by assigning each recording and transcription a participant ID. A coding frame was developed inductively, drawing upon themes from interviews. **How well was the data collection carried out?** Appropriately. There is great consideration for the 'participant fatigue' and sensitive nature of participants recalling childhood experiences of abuse. The research team provided in advance topics to be discussed with examples of sensitive questions that might be asked to ensure participants were aware of the interview process. In addition, the research team provided a trained	some bias into the sample. A small number of participants (n=13) were selected from the child maltreatment study to be followed up, while the remainder were recruited via an open, public invitation. Those who had previously taken part in the NSPCC child maltreatment study may have taken part in this follow up study because they had a previously established connection, and this may well have been the key to their participation' (p60). **Were the methods reliable?** Somewhat reliable. Although the study has two components - qualitative and quantitative findings, it is difficult to identify where the survey threads into the findings. **Are the data 'rich'?** Rich. The findings are rich and the context is provided about the participants i.e. sex, abuse suffered and length of time. The author states the number of young people who identify with each theme which makes useful comparisons throughout. **Is the analysis reliable?** Reliable. The process of analysis is well documented and the authors	and welfare, and the principle investigators have extensive sensitive interview experience. In addition, training was carried out with the research team by an expert consultant in domestic and sexual violence. The majority of interviews took place in participants' homes so the research team had an enhanced Criminal Records Bureau checks. Ethical approval was granted by the NSPCC Child Protection Research Ethics Committee. Participants gave informed consent and had the right to withdraw at any time. **Were service users involved in the study?** No. Service users did not co-produce the research. **Is there a clear focus on the guideline topic?** Yes. Q8: To explore impact of the Assessment Framework on practice. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes.	too, one part about bias and secondly, the purposive sample group was designed to target young people who had experienced systematic abuse. The findings do not provide a representative sample of the general young adult population nor those who have experienced abuse. As stated in limitations 'although the young people in this study did volunteer to discuss their childhood experiences, the researchers do not believe this means these young people represent a sample that is predisposed to disclose their abuse. If this were the case, the latency period for disclosure would be much shorter' (p60). **Overall assessment of external validity:** ++ Study relates to question of exploring young peoples' views and experiences of disclosure. The research design is rigorous and the authors' consideration of ethics is crucial to handling sensitive experiences effectively. **Overall validity rating:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
counsellor to support participants should they request it. Five participants in total accepted support.	code the transcripts to determine overarching themes. **Are the findings convincing?** Convincing. The findings are clearly presented and provide a wealth of information about the disclosure journey for a young adult reflecting upon their childhood experiences of abuse. Each theme is characterised with validating statements from research participants. **Are the conclusions adequate?** Adequate. The conclusions are supported by findings, so where an implication for practice is posed, the voice of the young adult and their experience is at the heart of that. For example, 12 young people reported that social workers were working with their family, but some professionals did not engage with them (n=4); ask the 'right question' (n=5); and acted insensitively when they did visit (n=2). Findings suggest the importance of talking to children and young people on their own.	Social work practitioners, managers, and staff from partner agency; also parents and children. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Participants were recruited from 5 sources, majority was from the NSPCC website and also through the maltreatment study. So participants had or were receiving support. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. The study relates to adult survivors experiences of disclosing abuse and seeks to explore key factors that promoted disclosure. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. The report contains a chapter (3) entitled 'disclosing abuse' that is dedicated to the young adults experience of disclosing abuse, both informally and formally. The findings reflect the adult survivors' viewpoints of their network and what helped and hindered disclosure.	A high quality study that meets the research aims and objectives and deals well with ethical concerns.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Does the study have a UK perspective? Yes. The study is informed by 60 young adults who reside in the UK.	

2. Beckett H, Brodie I, Factor F et al. (2013). 'It's wrong … but you get used to it' - A qualitative study of gang-associated sexual violence towards, and exploitation of, young people in England. London: Office of the Children's Commissioner for England

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. The research team conducts individual interviews with young people (n=150); 11 focus groups with professionals (n=76); and 8 single-sex focus groups (n=38). The comprehensive methods section details the rationale for interviewing participants because of the sensitive nature of the topic and to follow an ethical protocol. In addition, safeguarding concerns have been explored. **Is the study clear in what it seeks to do?** Clear. The forward from Sue Berelowitz, Chief Executive, Office of the Children's Commissioner details the context of the research: very little is known about the prevalence of sexual violence and exploitation within gangs by children and young people against other children and	**Is the context clearly described?** Clear. Under each direct quote, it is clear where data was collected, whether they are a young person or professional and age of participant (if individual interview). The individual interviews with young people (n=150) contain detailed characteristics, however the focus groups held with professionals (n=74) and young people (n=38), it is unclear on the characteristics of these participants. **Was the sampling carried out in an appropriate way?** Appropriate. Very clear that participants were recruited via agencies that were supporting young people to minimise risk. The authors state the potential for 'bias into the sample - and excludes other potential participants with valid contributions to offer - it was felt that the	**Does the study's research question match the review question?** Yes. The study explores 150 young people's and 76 professional's responses to gang-associated sexual violence and exploitation. The purpose is to understand the prevalence and experiences of young people: Chapter 4 is relevant to research question because it explores barriers that hinder young people formally disclosing. **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval was gained from four different research ethics committees and relevant local approvals were obtained within each research site. The research team was accountable to a research project advisory group, a young people's advisory group and local	**Overall assessment of internal validity:** ++ Very comprehensive, effective study with detailed findings that are summarised into recommendations. The methodology has dealt appropriately with the ethics and risk associated with the subject matter, and the research team have ensured the voice of the child is at the heart of the report. **Overall assessment of external validity:** ++ The study meets all criteria and has dealt effectively with ethical considerations. In addition, the interview guide was co-produced with the YPAG to make questions age appropriate. **Overall validity rating:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
young people. The purpose is to understand through interviews with young people and professionals' experiences to better inform national and local policy. **How defensible/rigorous is the research design/methodology?** Defensible. Very thorough research design and methodology which was governed and reviewed by a number of different bodies: Research Project Advisory Group; Young People's Advisory Group' and Site specific Multi-agency Advisory Groups. Qualitative interviews were conducted with 150 young people; 11 focus groups with 76 professionals; and 8 single-sex focus groups with 38 young people. There is a detailed breakdown of the 150 young people who participated in individual interviews, however the focus group held with professionals and young people is not descriptive. **How well was the data collection carried out?** Appropriately. Data collection section is thorough and the research team explained the measures to ensure the participants' comfort by facilitating the young people to talk in the third	risks of engaging those outside of services could not be adequately negated within a time-limited, large-scale, multi-site project such as this' (p12). **Were the methods reliable?** Somewhat reliable. The data is collected by one method, which were qualitative interviews. **Are the data 'rich'?** Rich. The research team cite references to where each finding was collected which helps contextualise responses to each participant. There are limitations as explored: 'Due to the flexibility built into the interviewing process, not all issues were covered with all of these interviewees' (p14). **Is the analysis reliable?** Reliable. Qualitative interviews were thematically analysed using NVivo 8 which underpin the findings in the research. The research team explain the executive decision to generally prioritise the young persons' voice to be presented in the report. **Are the conclusions adequate?** Adequate. The narrative findings	multi-agency advisory groups in each research site. **Were service users involved in the study?** Yes. In order to use age-appropriate research questions, the young people's advisory group co-produced the interview schedule. **Is there a clear focus on the guideline topic?** Yes. The relevant section is Chapter 4.2 where young people and professionals state factors that hinder disclosure: confusion about what actually constitutes sexual violence and exploitation; the acceptance of sexual violence and exploitation; and low levels of reporting and seeking support from professionals, i.e. judgement by others, lack of faith in services, perception of police and absence of conviction. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. 150 Young people's experience of gang-associated sexual violence and exploitation, and professionals (n=76) who have experience/specialism working with sexual violence and exploitation.	An excellent, thorough empirical study which meets its research aim and details implications for practice and policy on a local and national level.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
person, unless they wanted to actively choose otherwise, i.e. conversational manner using the interview schedule as a framework for discussion. There is effective consideration of the commitment to maintaining participants' confidentiality and anonymity. An ethical protocol was developed on the basis of 'no harm should come to any individual as a result of their agreement to facilitate or take part in the work' (p12).	of the voice and experience of participants contextualise the current knowledge and prevalence of sexual violence and exploitation in gangs. The relevant section to disclosure (Chapter 4) concludes that from the aim of 'identifying learning for embedding more effective systematic response to these issues in the future ... Prompted responses to these are now presented in the form of recommendations' (p51). The recommendations are structured to address national and local policy, which in the context of presenting findings from 6 different localities in England, map the issue with scope to respond.	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Young people were selected because they were/had received support from services, and professionals from statutory services were interviewed i.e. social care, police, and education. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Chapter 4 relates to barriers for young people disclosing sexual violence and exploitation to professionals. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. 'The research aimed to consider: the scale and nature of gang-associated sexual violence and exploitation in six areas of England; the main pathways into gang-related sexual violence and exploitation for young people living in these neighbourhoods; and potential models for an effective multi-agency response to the issue' (p6). **Does the study have a UK perspective?** Yes. Study is carried	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		out in 6 different research areas. For confidentiality purposes the sites are not named but do 'reflect a broad range of experiences of working with gangs and different demographic profiles' (p6).	

3. Burgess C, Daniel B, Scott J et al. (2012). Child neglect in 2011: an annual review by Action for Children in partnership with the University of Stirling. Watford: Action for Children

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. The authors conduct 12 focus groups with 114 professionals across 6 local authorities. The aim of the focus group was to 'gather more in-depth information about prevalence, recognition and response in relation to neglect' (p25). Participants were informed in advance as to topics of discussion at focus group. The author states that meetings were recorded and detailed notes were taken. There is no theoretical discussion as to the purpose of conducting focus groups. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?**	**Is the context clearly described?** Not sure. Little information. **Was the sampling carried out in an appropriate way?** Not sure. Not reported. **Were the methods reliable?** Reliable. Data is collected by two methods - focus group and surveys - and findings are justified within the data collected. **Are the data 'rich'?** Mixed. The annual review has stated that participants in focus group and survey represent a broad range of agencies, however without a breakdown of representatives, it is difficult to distinguish where the information came from.	**Does the study's research question match the review question?** Yes. The relevant section in the report is part 2 where data from professionals is gathered about how good services are at recognising children. **Has the study dealt appropriately with any ethical concerns?** No. Not reported. **Were service users involved in the study?** No. Service users did not co-produce this report. **Is there a clear focus on the guideline topic?** Yes. The annual review has a particular focus to what supports and hinders recognition of child neglect.	**Overall assessment of internal validity** - The annual review has carried out 12 focus groups which include 114 representatives from different agencies, however the findings and conclusions are 'somewhat convincing' because there is difficulty in identifying or contextualising who said what. There is no consideration of limitations or theory underpinning focus groups. **Overall assessment of external validity** ++ Overall, study meets most of the quality criteria however the study is not co-produced. **Overall validity score** -

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Somewhat defensible. The annual review is descriptive about the local authorities and participants (n=117) who were invited to focus groups, however a limitation is that there is not a clear sample method or demographic information so it is difficult to make generalisations. **How well was the data collection carried out?** Appropriately. A qualitative design is appropriate and authors are explicit that focus groups are recorded and detailed notes are taken.	At present, data appears anecdotal and there is no context to the narrative findings. **Is the analysis reliable?** Somewhat reliable. The project team states that data has been 'analysed in depth to look for emerging themes in the same way as the qualitative information from the survey above' (p26). The detail in the survey analysis is that open-ended questions were grouped under overarching themes and headings, with particular points of interest highlighted, as it was not possible to include everything. This process of analysis is not underpinned by theory or appear rigorous and could be subject to bias. **Are the findings convincing?** Somewhat convincing. Similar to the analysis, without a clear framework identified and the authors stating that they could not 'include all the detailed information' (p26), it is difficult to form a base of judgement on whether the findings are reliable. **Are the conclusions adequate?** Somewhat adequate. Again, the conclusions highlight overarching	**Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. The population of the annual review, in part, is based on professionals working across agencies experience and views of what helps and hinders recognition and identifying child neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Intervention is delivered across all agencies that come into contact with children, i.e. schools and children's social care. **Does the study relate to at least one of the activities covered by the guideline?** Yes. Relates to recognition. **For views questions) Are the views and experiences reported relevant to the guideline?** Yes. The aim is to 'gauge the current situation with regard to neglect and monitor the effects of changes in national and local policy' (p5).	The annual review meets the aim through the research design and mixed method data collection approach. The findings are representative of a large sample of professionals that work with children who are at the frontline for identifying and responding to child neglect. However, there is little information about consent of participants or what geographical region data is collected, so caution to generalise. Conclusions are difficult to see as reliable because the analysis is 'somewhat reliable'.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	themes but there is no consideration of limitations or clarity where data is from i.e. police officer or social worker, hence making conclusions difficult or reliable.	**Does the study have a UK perspective?** Yes. The review is carried out across local authorities within the UK.	

4. Children's Commissioner (2015) Protecting children from harm: A critical assessment of child sexual abuse in the family network in England and priorities for action. London: Office of the Children's Commissioner for England

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
1. Qualitative component 1 Call for evidence. **1.1 Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. **1.2 Is the process for analysing qualitative data relevant to address the research question?** Yes. All evidence analysed within overarching research framework. **1.3 Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. No contextualisation given to responses to call for evidence, e.g. which part of the country they came from, or what type of service. **1.4 Is appropriate consideration given to how findings relate to researchers' influence; for example,**	**4. Quantitative component 1** Survey of adult survivors of intrafamilial child sexual abuse. **4.1. Is the sampling strategy relevant to address the quantitative research question (quantitative aspect of the mixed-methods question)?** Unclear. Does not state how participants for the survey were recruited. **4.2 Is the sample representative of the population under study?** Unclear. There is no analysis of this in the report. **4.3. Are measurements appropriate (clear origin, or validity known, or standard instrument)?** N/A. **4.4. Is there an acceptable response rate (60% or above)?** Unclear. No response rate reported - unclear how many individuals were asked to complete the survey.	**a. Does the study's research question match the review question?** Partly. Overall aim of the study is to assess the scale and nature of child sexual abuse in the family environment in England. Some aspects of this are relevant to our review question. We have extracted data in relation to recognition and disclosure (sections 14.1, 14.2 and 14.3) and the impact of intervention (section 15.3), as these relate most closely to our review questions. **b. Has the study dealt appropriately with any ethical concerns?** Partly. The ethical approach for the study is outlined in Appendix A. However, this did not include getting ethical approval for the study. There is also no mention of how informed consent was obtained from participants. There is some description of how data will be stored.	**Overall assessment of internal validity** - **Overall assessment of external validity** + **Overall validity score** - Little methodological information provided, particularly regarding survey distribution, response rates and representativeness of resulting sample. Limited consideration of ethical issues in reporting.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
though their interactions with participants? N/A. **2. Qualitative component 2** Site visits. **2.1 Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. **2.2 Is the process for analysing qualitative data relevant to address the research question?** Yes. Standard pro forma used to collect and analyse data. **2.3 Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. Little contextualisation of any differences between sites. **2.4 Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?** No. **3. Qualitative component 3** Oral evidence hearings. **3.1 Are the sources of qualitative data (archives, documents, informants, observations) relevant**	**5.1. Is the mixed-methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and quantitative aspects of the mixed-methods question?** Partly. Mixed method design appropriate, but study does not make it clear what the relative contributions of different aspects were expected to be (e.g. oral evidence hearings compared to focus groups). **5.2. Is the integration of qualitative and quantitative data (or results) relevant to address the research question?** Yes. **5.3 Is appropriate consideration given to the limitations associated with this integration, such as the divergence of qualitative and quantitative data (or results)?** No. Little consideration of limitations of survey approach in general.	**c. Were service users involved in the study?** No. Service users involved as participants, but do not appear to have been involved in designing, conducting or interpreting study. No mention of service users on advisory panel of independent experts. **a. Is there a clear focus on the guideline topic?** Yes. **b. Is the study population the same as at least 1 of the groups covered by the guideline?** Partly. Views given by adult survivors and professionals working with sexually abused children and young people. **c. Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **d. Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study includes information relevant to recognition (Q6) and response (Q20). **f. (For views questions) Are the views and experiences reported relevant to the guideline?** Partly. Although important to note that, due to age of some of the survivors involved in the research, experiences of services may reflect past service arrangements and practice.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
to address the research question? Yes. **3.2 Is the process for analysing qualitative data relevant to address the research question?** Yes. Analysis according to themes identified elsewhere in the research. **3.3 Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. No consideration given to, e.g., differences in perspective between voluntary and statutory organisations. **3.4 Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?** No.		**g. Does the study have a UK perspective?** Yes. England.	

5. Cossar J, Brandon M, Bailey S et al. (2013) 'It takes a lot to build trust' - Recognition and Telling: Developing earlier routes to help for children and young people. London: Office of Children's Commissioner

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
1. Qualitative component 1 Content analysis of peer support website for young people, focusing on topics about abuse and neglect. **1.1 Are the sources of qualitative data (archives, documents,**	No quantitative components.	**a. Does the study's research question match the review question?** Yes. Study aims to examine young people's perceptions of abuse and neglect, and to explore their experiences of telling and getting help from both informal and formal sources.	**Overall assessment of internal validity** ++ **Overall assessment of external validity** ++ **Overall validity score** ++ Overall very good quality study

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
informants, observations) relevant to address the research question? Yes. Threads posted by young people experiencing or concerned about abuse and neglect. **1.2 Is the process for analysing qualitative data relevant to address the research question?** Yes. Sample accessed on a single day. Clear approach to coding thread content. **1.3 Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** Partly. Relatively little consideration of how representative or otherwise contributors to the forum may be. **1.4 Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?** Yes. Not applicable as analysis of online conversations between young people. **2. Qualitative component 2** Interview study.		**b. Has the study dealt appropriately with any ethical concerns?** Yes. NHS ethics approval was gained, and study also approved by University of East Anglia School of Social Work Ethics Committee. Particular ethical issues relating to each strand Content analysis of internet forum - the organisation running the forum gave their consent to the study. Individual posters on the site were not asked for consent, because the forum information is already publicly available. Interview study - Consent considered to be an ongoing issue, and young person able to stop interview at any time. Use of stop/go cards if the young person did not want to answer a question, or wanted to stop completely. Clarity about boundaries of confidentiality. Young people provided with details of further support. Focus groups - Consent obtained, including consent for audio-recording, and use of ground rules. Less consideration of what to do if participants became distressed or no longer wanted to take part. **c. Were service users involved in the study?**	with use of peer researchers, and good consideration of ethical issues. Good reporting of methods of analysis and clear findings.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
2.1 Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question? Yes. Interviews with young people aged 11–20. **2.2 Is the process for analysing qualitative data relevant to address the research question?** Yes. Detailed analysis of young people's experiences of recognition, telling and help. **2.3 Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** Partly. Some consideration of how age and type of abuse affected experiences, but not extensive. **2.4 Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?** Yes, e.g., consideration of different answers that young people gave in initial questionnaire compared to face to face interview. **3. Qualitative component 3** Focus groups with children and		Yes. A team of 6 young researchers to: refine research tools and materials, collect data, and analyse and disseminate findings. **d. Is there a clear focus on the guideline topic?** Yes. **e. Is the study population the same as at least 1 of the groups covered by the guideline?** Partly. The study involved multiple data collection strands. Most of these involved children and young people who had experienced or were at risk of abuse and neglect, their parents or carers, and professionals working with them. However, some of the young people (number not specified) in the interview study were adult survivors (aged between 18 and 20). Their data have been included here. **f. Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **g. Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to recognition	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
young people (data not reported), parents and professionals. **3.1 Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. **3.2 Is the process for analysing qualitative data relevant to address the research question?** Yes. **3.3 Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** Partly. **3.4 Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?** No.		and response (referred to as 'help' in the study'). **h. (For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **i. Does the study have a UK perspective?** Yes. England.	

6. Coy M. (2009). Moved around like bags of rubbish nobody wants: how multiple placement moves can make young women vulnerable to sexual exploitation. Child Abuse Review, 18, 254-266.

Internal validity - approach and sample.	Internal validity - performance and analysis.	External validity.	Overall validity rating.
Is a qualitative approach appropriate? Appropriate **Is the study clear in what it seeks to do?**	**Is the context clearly described?** Clear Characteristics of participants detailed in 'the sample of young women' - which evidences the age, type of	**Does the study's research question match the review question?** Partly Study explores young women's vulnerability to sexual exploitation if they	**Overall assessment of internal validity:** + **Overall assessment of external validity:**

Internal validity - approach and sample.	Internal validity - performance and analysis.	External validity.	Overall validity rating.
Clear Exploratory study to identify how local authority placement moves can impact on children and young women's vulnerability to selling sex. **How defensible/rigorous is the research design/methodology?** Defensible The author states that prior to the research she was an outreach project worker for women in the sex industry. They note the 'ethical dimension of balancing two roles was an integral part of the research design, but also an ongoing negotiation throughout the course of the research, based on prioritising women's welfare' (Coy, 2006; Nutt and Bell, 2002) (pg.257). Additionally, the methodology is based on feminist participatory action approach (combining narrative interviews with workshops as developed by Maggie O'Neill, 2011). **How well was the data collection carried out?** Not sure/inadequately reported Limited information on specificity of data collection, however author notes that the interviews were unstructured to enable the interviewees to narrate an account that 'includes those events that they frame as most significant in their lives, and retain control over the areas discussed' (p.256).	abuse, type of placement and memories of professional engagement. Throughout the text, the author provides anonymity to participants but details their age. **Was the sampling carried out in an appropriate way?** Somewhat appropriate Due to the authors dual role as outreach worker and researcher, interview participants approached the author to talk through their experiences. The author notes that 14 were included in the final sample, and 4 were not interviewed due to the ethical criteria because they were 'judged to be in need of primarily of professional support that would be compromised by inclusion in the interview' (pg. 257). **Were the methods reliable?** Somewhat reliable One form of data collection - qualitative study. **Is the analysis reliable?** Reliable All interviews are recorded and transcribed, these were then analysed using Doucet and Mauthner's (1998) 'Voice-Centred Relational Method' (this is where each section of the story is read several times to understand how the self is presented; for relationships with others; and finally	experience multiple care placement moves. **Has the study dealt appropriately with any ethical concerns?** Yes The study was granted ethical approval by the University of Staffordshire. Consent granted and participants able to withdraw at any time. Names are anonymised. **Is there a clear focus on the guideline topic?** Yes Selling sex under age of 18. **Is the study population the same as at least one of the groups covered by the guideline?** Yes Adult survivors of sexual exploitation. **Is the study setting the same as at least one of the settings covered by the guideline?** No Not reported. **Does the study relate to at least one of the activities covered by the guideline?** Yes Reflective accounts of young women who experienced sexual exploitation under age of 18.	+ Sound ethical consideration, however study only in part relevant to a research question (potentially recognition - risks). **Overall validity rating:** + Sound empirical study with thorough research method and justification. However, caution to generalise findings as represent 14 adult survivors experience of sexual exploitation in London. The potential for research bias, although described in the study, is not detailed in limitations.

Internal validity - approach and sample.	Internal validity - performance and analysis.	External validity.	Overall validity rating.
	for social and structural factors, pg. 258). **Are the findings convincing?** Convincing The author notes that the findings are a result of knowledge gained through their ethnographic involvement with women who sell sex as a pracitioner and a researcher. The themes have arose from a thorough analysis process (Doucet and Mauthner's 1998 'Voice-Centred Relational Method'). The findings do represent direct quotes from participants, that reflect their experience of local authority care placement moves, with vulnerability in sexual exploitation. **Are the conclusions adequate?** Somewhat adequate Author explores limitations initially i.e. 'small sample...not as a universal picture of cause and effect' (pg. 258). In the conclusion, suggests from supported findings that young women have found it difficult to develop trusting relationships with others and feeling unsettled due to frequent placement breakdowns. Unsure if the title is misleading.	**(For views questions) Are the views and experiences reported relevant to the guideline?** Yes Provide personal accounts of the impact of moving foster placements has on vulnerability of women. **Does the study have a UK perspective?** Yes London.	

7. Daniel B, Taylor J, Scott J (2010) Recognition of neglect and early response: overview of a systematic review of the literature. Child and Family Social Work 15: 248–57

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: The systematic review of literature examined an initial 20 480 items with an inclusion filtered to 63. The aim was to examine the three questions posed by authors: 1. What is known about the ways in which children and their families directly and indirectly signal their need for help? 2. To what extent are practitioners equipped to reorganise and respond to the indications that a child's needs are likely to be, or are being neglected, whatever the cause? 3. Does the evidence suggest that professional responses could be swifter? (p248). We have reported only the findings in relation to points 2 (recognition) and 3 (response), as the data reported in point 1 do not meet the evidence criteria for our related review question. **Appropriate and clearly focused question?** Yes. The systematic review aims to gather all relevant evidence relating to recognition of neglect and early response. The study narrows 20,480 papers to 63 with clear guidelines and criteria. The authors present where there are gaps in the evidence	**Inclusion of relevant individual studies?** Yes. The authors found that there were few studies that had been designed to fully meet the systematic review guideline and criteria. As stated, standardly systematic reviews include only randomised controlled trials (RCTs), however 'as the paucity of RCTs in this area, and the potential rich material in other types of studies, we included good-quality studies of all type' (p249). These were examined and ordered according to the standard hierarchy of evidence (Scharr 2006) in measures of confidence dependent on research design, methods and rigour. **Study quality assessed and reported?** Partly reported. With reference to relevant section: recognition, the authors present a narrative summary of relevant studies (n=5). 'The most direct evidence about the capacity of professionals to recognise neglect relates to health staff. Two-thirds of 513 staff in a children's hospital in Finland believed that they could recognise	**Does the study's research question match the review question?** Yes. The aim was to examine the three questions posed by authors: 1. What is known about the ways in which children and their families directly and indirectly signal their need for help? 2. To what extent are practitioners equipped to reorganise and respond to the indications that a child's needs are likely to be, or are being neglected, whatever the cause? 3. Does the evidence suggest that professional responses could be swifter? (p248). We have reported only the findings in relation to points 2 (recognition) and 3 (response), as the data reported in point 1 do not meet the evidence criteria for our related review question. **Has the study dealt appropriately with any ethical concerns?** Not reported. **Were service users involved in the study?** No. Services users did not contribute or co-produce the research.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + Thorough systematic review with clear criteria and guideline. However, the studies included are international. With the relevant section about recognition, the focus is different to guideline question. **Overall validity rating:** - Extensive systematic search, however little information given about individual included studies, and method for synthesising study findings very unclear.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
and where further research and policy changes are necessary. **Adequate description of methodology?** Partly adequate. Thorough methodology that presents process of systematic review and why the search is widened to incorporate other good quality studies with a different methodology. The authors are explicit about the criteria as to whether the study was included or excluded. However, relatively little information about method for synthesis. Little information given about individual included studies. **Rigorous literature search?** Yes. The systematic review explored 14 bibliographic databases.	maltreatment despite the associated difficulties' (Paavilainen et al. 2002). With this in mind, there is little evidence of critically appraising studies and no comparisons are made. This could be due to a lack of evidence found relevant to this section. **Do conclusions match findings?** Partly. New studies introduced in the Discussion/Conclusion section not reviewed in the earlier sections.	**Is there a clear focus on the guideline topic?** Yes. The systematic review conducts a thorough database search across 14 bibliographic to find material relating to neglected children and their parents who directly and indirectly signal for the help of professionals and how their needs are responded. The study has a focus on characteristics, the role of the professional, and the response. However, the study concludes that there is gaps in evidence and includes a dataset of 63. The area relevant to the topic is short and about who recognises child neglect and their capacity rather than what helps and hinders. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. The inclusion is of parents, children (aged pre-birth to 19) and professionals. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Cross-sectional context, i.e. hospital/education settings where professionals recognise and respond to child neglect.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Does the study relate to at least 1 of the activities covered by the guideline? Yes. The systematic review has 1 section entitled 'recognition' which discusses 5 studies that have explored professionals' role and effectiveness of recognising child neglect. **Does the study have a UK perspective?** Unclear. Some included studies from other countries.	

8. Gilligan P, Akhtar S (2006) Cultural Barriers to the Disclosure of Child Sexual Abuse in Asian Communities: Listening to What Women Say. British Journal of Social Work 36: 1361–77

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Mixed. The aims, objectives and methods are not very well stated - but you can find details about them woven into the narrative. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. It is clear how the discussion and conclusions have been reached, as they	**Is the context clearly described?** No. More detail is needed about the voluntary organisation from which participants from the community were invited to participate, and about the organisations from which the practitioners were recruited. **Was the sampling carried out in an appropriate way?** Not sure. Can't tell as there is a lack of adequate detail about the sampling strategy.	**Does the study's research question match the review question?** Partly. Includes views from practitioners on the process of recognising and assessing abuse and neglect, and on services providing early help for, or intervention following, abuse and neglect of children and young people but it doesn't include views of young people or adults who have experienced abuse or their carers. Rather it includes views of a community.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity rating:** - Research aims unclear, and study appears to bring in other sources of information such as practitioner data. Little information on sampling and methods.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
include references from the discussions with participants. However, it is not very clear throughout how to distinguish the practitioner voices from the community voices. **How well was the data collection carried out?** Somewhat appropriately. There is one very short paragraph on methods which doesn't provide enough detail about how participants were recruited, nor when in the timeframe of the study.	**Were the methods reliable?** Not sure. Can't judge because there is insufficient detail. **Are the data 'rich'?** Mixed. Only some information is relevant to this review. **Is the analysis reliable?** Not sure/not reported. There is a lack of detail about how the analysis was carried out. **Are the findings convincing?** Somewhat convincing. Conclusions and discussions are backed up by relevant transcript extracts but it is not always clear how community voices differ from practitioner voices. **Are the conclusions adequate?** Somewhat adequate. Conclusions are clear from what has been presented earlier but it would have been better for this review if practitioner voices had been more clearly distinguished from other community voices.	**Has the study dealt appropriately with any ethical concerns?** Partly. It mentions consent to report (anonymised) discussions was obtained from some of the participants and those are the ones the paper reports on. **Were service users involved in the study?** No. Young people with experience of abuse are not included, nor are their carers. **Is there a clear focus on the guideline topic?** Yes. The paper focuses on recognition and response to abuse and neglect which is of relevance to the review topic. **Is the study population the same as at least 1 of the groups covered by the guideline?** Partly Covers 1 area by including the views of practitioners working with children and young people who at risk of, are experiencing, or have experienced abuse and neglect.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Is the study setting the same as at least 1 of the settings covered by the guideline? Yes. Study is set in a community building - the scope settings include voluntary sector settings, including sports and youth clubs. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. The study covers: recognition of child abuse and neglect by practitioners working with children and young people. **(For views questions) Are the views and experiences reported relevant to the guideline?** Partly. Includes some practitioner views but difficult to disentangle them from other voices representing the community. **Does the study have a UK perspective?** Yes, Bradford.	

9. Harper Z, Scott S (2005) Meeting the needs of sexually exploited young people in London. London: Barnardo's

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. Unclear how practitioner participants were identified. **How well was the data collection carried out?** Somewhat appropriately. Interview protocol provided for practitioners but not young people. Unclear how interviews were recorded.	**Is the context clearly described?** Clear. **Was the sampling carried out in an appropriate way?** Not sure. Little information given regarding how individual participants were sampled. **Were the methods reliable?** Reliable. **Are the data 'rich'?** Mixed. Relatively good data from practitioner interviews, but analysis of input from young people relatively brief. **Is the analysis reliable?** Not sure/not reported. No mention of double coding of analysis. **Are the findings convincing?** Somewhat convincing. **Are the conclusions adequate?** Somewhat adequate.	**Does the study's research question match the review question?** Partly. Overall research questions do not match our review question, but contains some relevant information and has been included due to overall paucity of evidence on child sexual exploitation. **Has the study dealt appropriately with any ethical concerns?** Partly. Thorough ethical protocol covering informed consent, confidentiality, recording and storing data. However, no ethical approval sought. **Were service users involved in the study?** No. Service users involved as participants, but not in designing, conducting or interpreting study results. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children and young people at risk of or experiencing child sexual exploitation and professionals working with them.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** + Study was conducted in 2005, which means the findings may be somewhat outdated as awareness of, and practice in relation to, CSE has changed considerably since that time. Relatively sparse reporting of interviews with children and young people.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Is the study setting the same as at least 1 of the settings covered by the guideline? Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Contains information relevant to Recognition and Response. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes. England. However, important to note that study was conducted in 2005, since which time there has been much greater awareness of CSE, and significant changes to practice.	

10. Kazimirski A, Keogh P, Kumari V et al. (2009) Forced Marriage Prevalence and Service Response. London: Natcen

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Qualitative research forms part of a wider mixed methods study, but we have extracted data from qualitative element only.	**Is the context clearly described?** Not sure. An anonymised description of each case study local authority is given in the Methods section. When reporting findings, differences between the local authorities are reported, but not	**Does the study's research question match the review question?** Partly. Part of study is looking at prevalence of forced marriage (not relevant to review question) but part is looking at 'how services are currently responding to cases of	**Overall assessment of internal validity:** + **Overall assessment of external validity:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is the study clear in what it seeks to do? Clear. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. Good justification for selection of local authority case study areas. Less clear how individual participants were sampled and recruited. **How well was the data collection carried out?** Appropriately. Use of topic guide, and all interviews digitally recorded.	linked back to the initial description (e.g. referring to them as local authority A etc.). **Was the sampling carried out in an appropriate way?** Not sure. Little information given regarding sampling approach. **Were the methods reliable?** Somewhat reliable. Only interview data used - not triangulated with other sources of data. **Are the data 'rich'?** Rich. Good exploration of different perspectives, although no direct quotes from participants used in research. **Is the analysis reliable?** Reliable. Thematic analysis using specialist software, which allowed checking of extent to which interpretations of the data were shared across the research team. **Are the findings convincing?** Convincing. **Are the conclusions adequate?** Adequate.	forced marriage' (p11). There is content relevant to recognition, early help and response. **Has the study dealt appropriately with any ethical concerns?** No. No ethical approval gained. No description of how consent was gained from professionals involved in the research. Whilst there are fewer risks involved in interviewing professionals, some may have been directly affected by issues around forced marriage, so consideration of consent and support would have been beneficial. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Professionals working with young people at risk of, or experiencing, forced marriage. However, it should be noted that some of the professionals also worked with adults who were at risk of, or experiencing forced marriage. However, the majority of the report is	Good relevance to question, but no consideration of ethical issues. **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		concerned with practice in relation to children and young people. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study has content relevant to recognition, early help and response. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes.	

11. Liao LM, Elliott C, Ahmed F et al. (2013) Adult recall of childhood female genital cutting and perceptions of its effects: A pilot study for service improvement and research feasibility. Journal of Obstetrics and Gynaecology 33: 292–5.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate Aim of research is to explore women's recall of FGM and their perception of long-term consequences.	**Is the context clearly described?** Unclear Not clear where FGM was carried out. **Was the sampling carried out in an appropriate way?**	**Does the study's research question match the review question?** Partly Study relates to adult 'recall' of FGM rather than their views and experiences of recognition, assessment, early help, or response.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** -

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is the study clear in what it seeks to do? Clear **How defensible/rigorous is the research design/methodology?** Defensible Convenience sample appropriate for small pilot study. **How well was the data collection carried out?** Somewhat appropriately Unclear whether standardised interview schedule used, or how data were recorded.	Somewhat appropriate Appears to be convenience sample of women attending African Women's Clinic in London. **Were the methods reliable?** Not sure Unclear whether standardised interview schedule used. **Are the data 'rich'?** Poor Mostly quantitative analysis of the interview data. **Is the analysis reliable?** Somewhat reliable **Are the findings convincing?** Somewhat convincing **Are the conclusions adequate?** Adequate Within the research question set conclusions are adequate.	However, in view of the paucity of evidence we have found in relation to FGM, this study has been included. **Has the study dealt appropriately with any ethical concerns?** Partly Study approved as a service development project by University College London Hospitals NHS Foundation Trust Research and Development department. However, no mention in article of informed consent, informing participants of how information would be used and so on. **Were service users involved in the study?** No Service users involved as participants only. **Is there a clear focus on the guideline topic?** Yes **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes Adult survivors of FGM	**Overall validity rating:** - Lack of information regarding where FGM was conducted is a significant omission in terms of us being able to draw conclusions from this study relevant to our review.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Is the study setting the same as at least 1 of the settings covered by the guideline? Yes **Does the study relate to at least 1 of the activities covered by the guideline?** Partly Does not clearly map on to activities, but has been included due to paucity of evidence in this area. **(For views questions) Are the views and experiences reported relevant to the guideline?** Partly Again, does not map directly on to activities covered by the guideline but has been included due to paucity of evidence on FGM. **Does the study have a UK perspective?** Yes Although it is likely that at least some of the participants experienced FGM whilst living in other countries (this is not specified in the paper).	

12. McElvaney R, Greene S, Hogan D (2014) To tell or not to tell? Factors influencing young people's informal disclosures of child sexual abuse. Journal of Interpersonal Violence 29(5): 928–47

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Study seeks to explore in-depth experiences of young people and parents by directly understanding factors that influenced disclosure and how their experiences facilitated disclosure, as informed by other studies Jenson et al. (2005), and Staller and Nelson-Gardell (2005) who followed a similar methodology. Therefore authors have adopted a similar research design for interviewing 22 young people and 14 parents of these young people. This facilitated a thematic approach which found five key domains related to disclosure. **Is the study clear in what it seeks to do?** Clear. Clear research question relating to the 'factors influencing informal disclosure of child sexual abuse experiences, taking account of dynamics operating prior to, during and following disclosure' (p928). The study finds 5 common themes that influence disclosure from the perspective of children who have	**Is the context clearly described?** Clear. The authors provide information about the characteristics of the participants in a table which illustrates age and gender, and states where recruitment of participants for study happened. **Was the sampling carried out in an appropriate way?** Not sure. The authors provide little information how sampling was carried out albeit, it is known that participants had been assessed and were deemed credible in their disclosure of child sexual abuse by a professional working within a hospital setting. **Were the methods reliable?** Somewhat reliable. The data is collected by 1 method, which were qualitative interviews. Comparisons are made between other studies findings and provides a basis and discussion for the research piece with what is known about children disclosing abuse.	**Does the study's research question match the review question?** Yes. The study is in relation to the review question of gathering young people's views and experience of recognition and disclosure. **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval was obtained from both the hospital's ethics committee where participants were recruited for the study, as well as the university's School of Psychology ethics committee. Parents consented to be interviewed in writing and assent from young people were obtained from all participants in this study. **Were service users involved in the study?** No. Service users have not co-researched this paper. **Is there a clear focus on the guideline topic?** Yes. The study has a focus in relation	**Overall assessment of internal validity:** + The study explores the justification of adopting the methodological technique however there is little consideration for the limitations and generalisability of the study findings. **Overall assessment of external validity:** ++ Overall, study meets most of the quality criteria however caution to generalise the UK as the study is based in a child sexual abuse assessment and therapy centre in a hospital in Ireland. **Overall validity rating:** ++ A good, thorough empirical study which meets its research aim.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
experienced child sexual abuse and their parents. The conclusions draw upon the complex intrapersonal and interpersonal dynamics that factor into the disclosure process. The literature is appropriate to set the current knowledge and recognise studies that have similar methodologies. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. The study provides a breakdown of the participants although the predominance of female interviewees is a possible limitation. The author is clear where the participants were recruited - a hospital setting where children were accessing support or had been assessed for child sexual abuse, although no further information is provided in the way in which recruitment was carried out. There is little demographic information. **How well was the data collection carried out?** Appropriately. The author is explicit in the theory adopted to inform methodology - Grounded theory, in addition providing a wealth of information in the interview schedule, i.e.	**Are the data 'rich'?** Rich. The data is well founded in grounded theory and was data managed and analysed in NVivo. The authors give justification with other studies (Charmaz 2006; Strauss & Corbin 1998). **Is the analysis reliable?** Reliable. The coding process is thorough where McElvaney et al. (2012) inform the 'triangular model' that 'describes the analytic process as moving from the raw data transcripts toward higher level conceptual categories and domains'. The research team use analytic memos and describe the theme driven approach to reflective an active process of coding. Themes are verified by research participants - a young person and parent - to provide credibility checks by reading their transcript and going through codes with a member of research team. The thorough methodology highlights the accuracy and depth of the data collecting process. **Are the findings convincing?** Convincing. The findings are clearly presented	to guideline topic because the findings explore young people and parents' perspective on factors that affect disclosure and consequently, recognition. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. The population of the study is based on individual qualitative interviews with parents and young people who were accessed by the research team in a child sexual abuse assessment and therapy service, based in a children's hospital in Ireland. 'All child participants had given an account of sexual abuse that was deemed credible by professionals who assessed them' (p932). **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Intervention is delivered in a children's hospital in Ireland. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Relates to recognition.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
semi-structured with open questions to elicit information accordingly. Authors are explicit that interviews were recorded subject to consent of interviewee.	and are supported by extracts from data collection. Comparisons are made and author meets the research aim in a coherent way illuminating 5 inductive domains found by children who have experienced child sexual abuse and their parents which influence the process of disclosure: being believed, being asked, shame/self-blame, fears and concerns for self and others, and peer influence. **Are the conclusions adequate?** Adequate. The conclusions are well founded in the context of the paper with implications for practice - the findings of this study suggest that 'many factors combine to influence a child's readiness and ability to tell' (p944). The paper meets the aim by threading intrapersonal and interpersonal factors that impact on the disclosure process. The author provides solutions for professionals to consider (with a growing body of evidence from other studies) to suggest that peers provide a good support network - thus, educating young people with the disclosure process could help.	**(For views questions) Are the views and experiences reported relevant to the guideline?** Yes The aim explores factors that influence informal disclosure of child sexual abuse experiences and considers the complex intrapersonal and interpersonal dynamics reflecting the conflict inherent in the disclosure process. **Does the study have a UK perspective?** No. The study is based in Ireland, therefore caution to other settings.	

13. McNaughton Nicholls C, Harvey S, Paskell C (2014) Gendered perceptions: what professionals say about the sexual exploitation of boys and young men in the UK. London: Barnardo's

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Study notes that 'qualitative research enables an in-depth exploration of social phenomena and practices, and is particularly suited to exploring emerging and complex issues' (Lewis and McNaughton Nicholls 2014) (p13). **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. Purposive sampling of interviews to represent a range of regions of England, types of service, seniority, length of service and gender. **How well was the data collection carried out?** Somewhat appropriately. Little information given regarding data collection. Some participants were 'interviewed online' (p15) - unclear how this was conducted,	**Is the context clearly described?** Clear. **Was the sampling carried out in an appropriate way?** Appropriate. Purposive sampling to obtain a spread of geographical location, professional settings, gender and levels of experience. **Were the methods reliable?** Somewhat reliable. Little detail regarding how online interviews were conducted. **Are the data 'rich'?** Mixed. Little consideration of divergence in perspectives along lines of geography, professional background and so on. **Is the analysis reliable?** Somewhat reliable. Analysis conducted using NVivo software. Unclear if procedures such as double-coding of interviews was used.	**Does the study's research question match the review question?** Partly. The study has 4 research questions, 1 of which matches our review question which is - to 'suggest ways in which policy and practice may be able to identify and appropriately respond to male victims of CSE, as well as those at risk' (p13). The other 3 questions are less relevant to this review question which are - identify perpetration and victimisation processes apparent in male-victim CSE cases known to professionals - explore existing service provision for boys and young men at risk of or experiencing CSE - identify future research priorities (p13). **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval by NatCen's research ethics committee and clarity regarding how data would be presented and stored.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + UK study but only part of overall research aim was relevant to our review question. **Overall validity rating:** + Only part of overall research aim was relevant to our review question. Study is of reasonable quality, although limited exploration of divergent perspectives across different types of interviewees.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
and no analysis of the impact of this on data.	**Are the findings convincing?** Convincing. **Are the conclusions adequate?** Adequate.	**Were service users involved in the study?** No. The report makes occasional references to a young people's workshop. However, this is not described in the methods section. Any findings reported from this strand have therefore not been extracted here. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Professionals working with boys and young men experiencing, or at risk of, sexual exploitation. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Relates to recognition (and response).	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		(For views questions) Are the views and experiences reported relevant to the guideline? Yes. **Does the study have a UK perspective?** Yes, England.	

14. NSPCC (2013) Would they actually have believed me? A focus group exploration of the underreporting of crimes by Jimmy Savile. London: NSPCC

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Study seeks to explore experiences - qualitative approach therefore appropriate. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. Original design (5 focus groups with a total of 50 participants) would have been more robust. However, drop-out from the study meant that there were only 26 participants in total. Also unclear why	**Is the context clearly described?** Unclear. Little information on the participants, or the context of their experiences. **Was the sampling carried out in an appropriate way?** Somewhat appropriate. Sample comprised those who had come forward to Operation Yewtree, and were prepared to take part in a focus group. There is no consideration in the report about how this may have affected people's views. **Were the methods reliable?** Somewhat reliable.	**Does the study's research question match the review question?** Yes. Study is about what helps and hinders children and young people from disclosing abuse by seeking information from victims of Jimmy Savile about what had prevented them from reporting to the police at the time of the abuse, and 'to explore how police can improve their management of the reporting process and subsequent interviews and contacts' (p4). **Has the study dealt appropriately with any ethical concerns?** No. No mention of ethical considerations within the report. This is	**Overall assessment of internal validity:** - **Overall assessment of external validity:** - Lack of consideration of ethical issues, and of transferability of findings given the very particular circumstances involving a high profile celebrity and subsequent documentary film raising awareness. **Overall validity rating:** - No consideration of ethical issues reported. Little consideration of transferability of the findings to other cases of abuse, given the particular circumstances (i.e. high

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
focus groups rather than individual interviews were chosen. **How well was the data collection carried out?** Somewhat appropriately. Broad questions asked of the groups are given in the report. Groups were recorded and transcribed.	**Are the data 'rich'?** Poor. Little exploration of diversity of views within the focus group samples. **Is the analysis reliable?** Somewhat reliable. Data have been analysed per question. It is unclear how this was undertaken. **Are the findings convincing?** Somewhat convincing. Given the lack of information regarding the study sample, and relatively little exploration of diversity in the data, difficult to have strong confidence in the findings. **Are the conclusions adequate?** Somewhat adequate. Given the lack of information regarding the study sample, and relatively little exploration of diversity in the data, difficult to have strong confidence in the conclusions.	quite a serious omission given the nature of the subject being discussed. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Partly. Four of the 26 participants were abused as adults rather than in childhood. Also, it should be noted that the individuals in this study had experienced abuse by a high profile person, which may mean that not all aspects of their experience are transferable to other experiences of abuse. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes.	profile celebrity case). Little detail given regarding participants, or methods of analysis.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Study relates to recognition (disclosure). **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes.	

15. Pearce J, Hynes P, Bovarnick S (2009) Breaking the wall of silence: practitioners' responses to trafficked children and young people. London: NSPCC

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Defensible. Good rationale given for why children and young people not directly involved. Substantial sample size. **How well was the data collection carried out?** Appropriately. Focus groups and interviews guided by a specific set of topics	**Is the context clearly described?** Clear. Contextual data provided for the three research sites. **Was the sampling carried out in an appropriate way?** Somewhat appropriate. Purposive sampling used for interviews. Practitioner focus groups and case files appear to be convenience sampled, characteristics of both samples well described. **Were the methods reliable?** Reliable.	**Does the study's research question match the review question?** Partly. Study has a range of research questions. The following were judged to be relevant to our review questions: Recognition: 2. Explore the obstacles that might emerge to identifying the numbers of young people trafficked in the three areas. Response: 7. Identify how the professionals feel these needs are best met. 8. Where possible, identify perceptions of how the children/young people feel these needs are best met. We considered questions 2 to be relevant to our review question on	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** ++ **Overall validity rating:** ++ Thorough data collection, analysis and reporting.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
and digitally recorded. Data recorded from case files using an agreed template.	**Are the data 'rich'?** Rich. Very detailed analysis, drawing out good distinctions between UK children who are trafficked and children trafficked from abroad. **Is the analysis reliable?** Reliable. Focus group data analysed by 2 members of research team. Thematic analysis cross-checked between two staff members. Interview data analysed using NVivo. Case files do not appear to have been double coded. **Are the findings convincing?** Convincing. **Are the conclusions adequate?** Adequate.	Recognition, and 7 and 8 to be relevant to our review question on Response. **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval for the research project was given by the University of Bedfordshire, School of Applied Social Studies Ethics Committee and by the NSPCC Ethics Committee (p47). Carried out in accordance with ESRC and British Sociological Association guidelines and Barnardo's Research ethics. **Were service users involved in the study?** No. Children and young people were not directly involved - information gathered via analysis of case files. This was due in part to ethical issues associated with involving them in the research. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Trafficked children and practitioners working with them.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Is the study setting the same as at least 1 of the settings covered by the guideline? Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. There is material relating to Recognition and Response. **Does the study have a UK perspective?** Yes. England.	

16. Rees G, Gorin S, Jobe A et al. (2010). Safeguarding young people: Responding to young people 11 to 17 who are maltreated. London: The Children's Society

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Study seeks to explore young people's views and experiences. **Is the study clear in what it seeks to do?** Clear. Clear research question relating to young people's views on seeking and receiving help. The aim is to better understand the experiences of young people in order to better meet their needs and improve the safeguarding system. Implications	**Is the context clearly described?** Clear. Characteristics of young people are described thoroughly and the author provides contextual information collected at interviews about the maltreatment. There is a balanced ratio of female/male participants involved in the study (10:14), who also represent a diverse locality, background and ethnicity. The limitations of conducting interviews with young people are discussed in depth where the authors give a clear explana-	**Does the study's research question match the review question?** Yes. The aim of the study is to explore 'access to, and initial responses of, services for young people with potential maltreatment … to promote protective responses for this target group' (p7). The section relevant to this review question is entitled young peoples' experience of seeking help. **Has the study dealt appropriately with any ethical concerns?** Yes.	**Overall assessment of internal validity:** + (When taking into account additional info from Jobe and Gorin.) The study does not have a rigorous methodology or consideration of limitations. Presentation of information is difficult to ascertain where data is collected making it challenging to draw conclusions. In addition, there is discrepancy in young people's age as referred in text to both: 11–17; and 11–18. **Overall assessment of external validity:**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
for practice are discussed in conclusion. **How defensible/rigorous is the research design/methodology?** Defensible. A qualitative design is appropriate given their research question. **How well was the data collection carried out?** Appropriately. The study says that interviews were recorded and transcribed. Study reports that interviews were carried out face to face, that they asked young people about seeking help, being referred to CSC and subsequent responses, and recorded with young person's consent.	tion for their decision-making process, although their sampling techniques are not clearly described. There is consideration of bias: 'our findings may arguably be a partial representation of events as we are unable to present the views or recollections of any of the professionals young people refer to' (p432). **Was the sampling carried out in an appropriate way?** Somewhat appropriate. Sample appears to be a convenience sample, rather than aiming to be representative of particular categories, however there is diversity of age, gender and ethnic background within the sample. Unclear how young people were selected for interview. **Were the methods reliable?** Somewhat reliable. The data was collected by 1 method which was qualitative interviews. **Are the data 'rich'?** Mixed. The authors state they have included participants that represent a different locality, background and ethnicity, however there is no	The research had ethical approval from the Institute for Research in the Social Sciences Ethics Committee, University of York and the Association of the Directors of Children's Services. **Were service users involved in the study?** No. Service users did not contribute or co-produce the research. **Is there a clear focus on the guideline topic?** Yes. The population of the study is qualitative interviews with young people who discuss their experience of disclosure. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. The participants of the study are 11–17 year old young people who are experiencing statutory support for maltreatment. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Local authority.	++ Study relates to question of exploring young peoples' views and experiences of recognition. **Overall validity rating:** + The study is suitable for scope and the findings enrich discussion about barriers to young people disclosing sexual abuse. Drawing on additional information from Jobe and Gorin (2013), where the research design is more informed, the findings are more convincing as data is richer and analysis is clearer.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	recognition of how different young people access a service i.e. the study includes 5 unaccompanied asylum seekers with little recognition that they might have a different experience of disclosing. The narrative findings of the young people make it difficult to distinguish where the information came from. The author does not state the number of young people that experience what and how so difficult to contextualise and responses are not compared or contrasted across groups. **Is the analysis reliable?** Somewhat reliable. Authors report that they used NVivo and have done a thematic analysis. **Are the findings convincing?** Somewhat convincing. There is relatively little presentation of the data analysis on which to base a judgement of whether the findings are reliable. There is also little consideration of diversity in views, for example the experiences of children who had sought asylum. **Are the conclusions adequate?** Somewhat adequate.	**Does the study relate to at least 1 of the activities covered by the guideline?** Yes. The aim explores 24 young people (11–17) who have been referred to Children's Social Care Services in England and have received statutory support. The paper has a particular focus on the young peoples' experience of disclosing and seeking help for maltreatment. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. The study includes one section entitled 'young people's experiences of seeking help'. Interviews have contributed to the findings which are reported to be divided into four categories: the difficulties with seeking help; seeking help from peers; seeking help from family members; seeking help from professionals. The other sections of the paper are not relevant to the current review question. **Does the study have a UK perspective?** Yes. Young people are accessed from	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	The conclusions draw out over-arching themes, without consider-ing diverse experiences within the group, for example the children who had sought asylum.	six English local authority areas and represented a range of ethnic backgrounds and ages.	

17. Rigby P (2011) Separated and trafficked children: The challenges for child protection professionals. Child Abuse Review 20: 324–40

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Study seeks to understand challenges in practice. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Defensible. **How well was the data collection carried out?** Somewhat appropriately. Lack of clarity regarding interviewees compared to focus groups, and which topics were discussed via which method.	**Is the context clearly described?** Unclear. Little description of context of child trafficking in Glasgow, e.g. the extent of the issue, whether this is a relatively new phenomenon etc. **Was the sampling carried out in an appropriate way?** Not sure. Method for selecting case files and interviewees not reported. **Were the methods reliable?** Reliable. **Are the data 'rich'?** Not sure. Little presentation of data within the report. **Is the analysis reliable?** Somewhat reliable. Analysis method described as 'broadly grounded theory'. Analysis does not appear to have been cross-validated by a second researcher or similar. **Are the findings convincing?** Somewhat convincing. Little presentation of primary data	**Does the study's research question match the review question?** Partly. Main research question is about identifying challenges emerging for practitioners working with separated children who have been trafficked. However, there is a section on 'Identification and assessment'. **Has the study dealt appropriately with any ethical concerns?** Partly. Author states that research governance provided by local Child Protection Committee but does not appear that any formal research ethics approval sought. Does not appear that consent was obtained from children for using their case files for analysis purposes. **Were service users involved in the study?** No. Research not co-produced by service users. **Is there a clear focus on the guideline topic?** Yes.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + Assessment is a subset of overall research question. **Overall validity rating:** - Not clear which data were gathered via interview, and which via focus group. Analysis methods unclear. Relatively little reference to, or presentation of, primary data gathered.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	to back up points made in analysis. **Are the conclusions adequate?** Adequate.	**Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Population is children who have been trafficked and the professionals working with them. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to assessment. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes, Scotland (Glasgow).	

18. Stanley N, Miller P, Foster Helen R (2012) Engaging with children's and parents' perspectives on domestic violence. Child and Family Social Work 17: 192–201

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. The study explores motivations for conducting focus groups with young people and the limitation/risk associated with 'promoting consensus and socially acceptable attitudes' (Kitzinger 1995). In-depth interviews were held with parents using a semi-structured interview schedule. In total, 19 young participants contributed in 5 focus groups; 11 survivors; and 10 perpetrators perspectives were explored to enrich the knowledge base of lived experience of domestic violence and professionals they have worked with. Therefore a qualitative approach provides an abundance of information to well-founded findings. **Is the study clear in what it seeks to do?** Clear. The research is clear that it seeks to identify 'the need for practitioners to engage with the emotional content of disclosure of domestic violence and to undertake this work in separate sessions with	**Is the context clearly described?** Unclear. No information about the context and little consideration about the disparity across participants from two different localities across the UK. **Was the sampling carried out in an appropriate way?** Not sure. There is no information about sampling. **Were the methods reliable?** Somewhat reliable. The data was collected by one method - qualitative interviews and focus groups, and no comparisons are made. **Are the data 'rich'?** Mixed. Data is founded in ground theory principles, therefore themes have been inductive however the authors omit that the clusters are down to interview schedule guide structure, which was not co-produced. The findings illustrate these themes with various quotes	**Does the study's research question match the review question?** Partly. In part, the study is relevant to the question because there are findings relating to disclosure and acknowledging domestic violence, and listening to and validating accounts, namely professionals. There is little information about the recognition of DV and the study is more about intervention. **Has the study dealt appropriately with any ethical concerns?** Partly. Ethical approval was sought and awarded by the University of Central Lancashire, however no information about ethical approval from local authority. Authors are explicit that consent was sought from all research participants. **Were service users involved in the study?** No. Service users did not co-produce the research. **Is there a clear focus on the guideline topic?**	**Overall assessment of internal validity:** + **Overall assessment of external validity:** - The study has relevance in part to exploring young peoples' perspectives on domestic violence, yet as study has adult perspective too, the voice of the child is represented generically. There is no information about obtaining ethical approval from local authorities. The authors are explicit that the research participants are anonymised and consent was sought. **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
parents and with children so that differing accounts can be heard safely' (p129). The evidence presented in the findings and discussion clearly illustrate experiences of young people and parents who require professionals to be 'perceived as powerful and effective in their response to domestic violence'. **How defensible/rigorous is the research design/methodology?** Defensible. Methodology section explores the purpose of focus groups with young people in order to create a familiar environment where the young people are used to attending and discussing abuse. The authors acknowledge limitation and risk through other studies (Kitzinger 1995). There is a theoretical basis underpinning analysing data using grounded theory principles (Strauss & Corbin 1990; Richie and Spencer 1994). In addition, assigning descriptive and analytic codes to explore 'commonalities and differences between the three groups were noted' (p194). No information regarding sampling of participants - but assumption is there were volunteers to partake	from participants to validate, however unclear on numbers of participants that agree with statements made. Discussion focuses upon implication for practice founded in other studies and participant comments, e.g.: 'All three groups of participants - young people, survivors and perpetrators - echoed the messages of other research with children and parents in their need for practitioners to listen to and validate their accounts: when stigmatized behaviour is exposed to external scrutiny, individual stories need to be heard and treated with respect' (p198). **Is the analysis reliable?** Somewhat reliable. Analysis is not reported beyond stored in NVivo, and founded in ground theory principles. There is an agreement amongst researchers for themes to be represented in findings. **Are the findings convincing?** Somewhat convincing. The study does meet the aim to explore children and parents' perspectives on domestic violence and there is a good representation of participants. Findings are clearly presented and supported	Partly. There is partly relevance to identified barriers and support in recognition, however the study explores experiences, therefore there is a focus on intervention. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Three groups of participants who have experienced domestic violence and are receiving support from statutory and voluntary services: young people; survivors; and perpetrators. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Interventions are delivered in statutory and voluntary services. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. The study is in relation to children and parents experiencing domestic abuse. **(For views questions) Are the views and experiences reported**	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
who were already accessing domestic violence services, both statutory and voluntary. **How well was the data collection carried out?** Somewhat appropriately. Little detail regarding data collection, only that interviews and focus groups adopted a similar semi-structure approach to facilitate exploration of domestic violence experiences of participants and the professionals that supported them. Data is stored and collected on NVivo. The authors are explicit that data is recorded and transcribed subject to consent of participant, however no information on whether any participant withdrew.	by extracts from research participants, however there is a feel of the evidence being manufactured to support a discourse alluded to in the literature review that domestic violence is prevalent in disadvantaged communities, among low income families and single children and that families using social services are likely to be this populace (p192). Therefore, experiences are representative of 'disadvantaged communities' perspectives and the author does not explore socioeconomic demographics within limitations. In addition, limitations are not explored nor is there recognition that different localities have different protocols for domestic violence. The young peoples' voice does not echo throughout and there is little depth beyond statements (see p197). **Are the conclusions adequate?** Somewhat adequate.	**relevant to the guideline?** Yes. The study explores experiences of survivors and perpetrators of domestic violence through qualitative interviews conducted with young people and parents. There is little information about the recognition of maltreatment of young people, as this is a cross-sectional study that explores both adults and children. **Does the study have a UK perspective?** Yes. The study is undertaken across two local authorities in the North and South of England.	

19. Tucker S (2011) Listening and believing: an examination of young people's perceptions of why they are not believed by professionals when they report abuse and neglect. Children and Society 25: 458–69

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. The study includes a thorough methodology to determine the purpose of conducting 102 qualitative interviews. The use of qualitative interviews and an advisory group is contextualised and supported by other studies (Roberts 2004). Authors considered risk and provided additional support to young people if they were distressed. **Is the study clear in what it seeks to do?** Clear. Clear research question relating to young people's views on seeking and receiving help. **How defensible/rigorous is the research design/methodology?** Defensible. Thorough and descriptive, supported through other literature. **How well was the data collection carried out?** Appropriately. Informed consent was given in an appropriate written form (Cohen et	**Is the context clearly described?** Clear. Characteristics are presented in terms of age and gender, as defined in table 1 (p60). The author does not include details of ethnicity. Additionally, the author notes the matter than all participants were self-selecting making it impossible to control for gender and age. Participants are appropriately provided anonymity. **Was the sampling carried out in an appropriate way?** Appropriate. Defined as a system of 'snowball sampling' (Cohen et al. 2004) which was adopted where earlier recruited participants were invited to encourage others to come forward and tell their stories. Author states strategy was slow but ensured strong commitment and participation. **Were the methods reliable?** Somewhat reliable. The data was collected by one method which were qualitative interviews.	**Does the study's research question match the review question?** Yes. Young peoples' perception of factors hindering recognition, i.e. disclosure. **Has the study dealt appropriately with any ethical concerns?** Yes. Sound ethical consideration. Ethical scrutiny was granted through researcher's institutional Ethics Committee. **Were service users involved in the study?** Yes. Study is co-produced through a self-selective advisory group consisting of 8 volunteers (3 females and 5 males) who participated in the study. A written brief was provided for all participants via email and this was followed up by an individual phone call. The group liaised via telephone conference style calls and email exchanges. **Is there a clear focus on the guideline topic?** Yes.	**Overall assessment of internal validity:** ++ Good qualitative study with large sample (n=102) and through methodological approach. **Overall assessment of external validity:** ++ Study relates to question of exploring young peoples' views and experiences of disclosure. Sound ethical consideration and the study is co-produced by an advisory group. **Overall validity rating:** ++ Good empirical qualitative study with large sample (n=102) and through methodological approach. The research is relevant to inform young peoples' perception of disclosure.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
al. 2004). Additionally, the research team offered a variety of methods of participation to enable geographical constraint to not be an issue.	**Are the data 'rich'?** Rich. Inferences from various conversations were only drawn on the basis of the frequency of their revelation and these were then used to generate specific categories, which described key factors and issues likely to contribute towards young people not being believed when they report abuse and neglect. This is stated to ensure data is representative of a range of views expressed by the young people (pp459–60). **Is the analysis reliable?** Reliable. The emerging typology was constructed directly from the data produced through the interviews. Interviews were transcribed, analysed and coded into identify common themes (Cohen et al. 2004). **Are the findings convincing?** Convincing. Findings are supported by an emerging typology - as directed by qualitative interviews and advisory group – generally represent the circumstances of young people disclosing to a professional. Figure 1 (p463) presents consideration of the typology of disbelief	The population of the study is qualitative interviews with young people who discuss their experience of disclosure. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Adult survivors (age 18–23) who had attempted to disclose abuse or neglect on more than one occasion to a professional who played a significant role in their life. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Child abuse and neglect. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Disclosure. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. Interviews have contributed to findings that are reported under the construction of the 'typology of	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	and the percentage reporting under each specific circumstance: background and baggage; family matters; reluctance and refusal; and personal relationships. Interwoven are direct quotes from interviews with male and females. Additionally, the advisory group provides a concise, succinct synthesis of each category. Literature supports the findings (Griffin 1993; Howarth 2007; Young 2006; and others). **Are the conclusions adequate?** Adequate. Implications for practice and policy are considered to suggest that support for young people at the time of disclosure could be significantly improved. Limitations are explored: 'Those involved in the advisory group in particular recognised this was 'their take' on the situations that had confronted a number of young people. They also recognised that practitioners might explain the issues and dilemmas in different ways. Yet, at the same time, it was argued that 'when others try and explain matters away from where they sit how often are young people really involved in trying to understand why?' (p468).	disbelief'. Related factors and issues are brought together through the development of specific 'circumstance categories': background and baggage; family matters; reluctance and refusal; and personal relationships (p462). **Does the study have a UK perspective?** Yes, England.	

Review question 6 – Findings tables

1. Allnock D, Miller P (2013) No one noticed, no one heard: a study of disclosures of childhood abuse. London: NSPCC

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: Researchers interviewed 60 young adults (aged 18-24 years) who had experienced high levels of different types of abuse and violence during childhood. The young adults were asked whether they had tried to tell anyone about what was happening to them, and what had happened as a result of their disclosures. Data suggests that disclosure was delayed from the start of the abuse by an average age of 7.8 years. This report describes their childhood experiences of abuse. It looks at whether they disclosed their abuse: what prevented them from disclosing, and the key factors that promoted disclosure. Conclusions highlighted 3 key themes: information; communication; and	**Participants** Adult survivors of child abuse - Young adults were aged 18 to 24. **Sample characteristics** • Age - The participants of the study are between ages of 18–24 (at time of interview in 2009/10). • Sex - The sample includes 7 males and 53 females. • Ethnicity - The sample were predominantly White (92%), with representation of Black or Minority (8%). • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - 95% of participants (n=57) of the young people reported contact sexual abuse and 44 experienced sexual abuse with an average length of abuse calculated at 7.8 years. The study also explores polyvictimisation - where 82% (n=49) reported 15 or more positive responses to different maltreatment and victimisation types in their lifetime. The authors	**Narrative findings** The relevant findings of this study are categorised into 4 sections and can be found in Chapter 3 (p.16-48): (1) What prevented disclosures in childhood; (2) what promoted disclosures; (3) informal disclosure through mother and/or friends; (4) formal disclosure which shines a spotlight on teachers, social services and the criminal justice system. 'Each section examines disclosure themes for two groups of young people – those who experienced sexual abuse (n=44) and those who experienced other forms of abuse (n=16). 1. What prevented disclosure in childhood: The barriers identified in the study for seeking help and why there was a long period before disclosure, reported in the study on average as 7.8 years (p18), is that: (1.1) young people had no one to turn to; (1.2) perpetrator tactics; (1.3) developmental barriers; (1.4) emotional barriers and anxieties; (1.5) no one listened/asked; and (1.6) anxiety over the confidentiality of their information (p24). 1.1. 'Young people had no one to turn to': 45% (n=27) of young people described feelings of isolation in childhood which curtailed their disclosure due to poor family relationships and adverse family circumstances. Isolation as a common theme is described to be geographical, physical and psychological. Some young people felt isolated by their experience, as stated 'I did speak to my friends at school and stuff,	**Overall assessment of internal validity:** ++ All of the criteria has been fulfilled to a high standard. The authors met all ethical concerns and sought solutions for young adults who may be affected by their participation in the study. There is a large sample group (n=60) and their characteristics are clearly described in the report. Findings prompt implications for practice so as to tailor better support and intervention for childhood abuse. The limitations are defined too, one part about bias and secondly, the purposive sample group was designed to target young people who had experienced systematic abuse. The findings do

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
being noticed, asked and heard. There are implications for practice. **Methodology:** The study used a mixed method approach to collect data from 60 young adults aged 18–24. For the purpose of findings relevant to the review question, data has only been extracted from qualitative methods, and so only this element of the research has been critically appraised. Qualitative data was collected via in-depth interviews. The interviewees were selected after completion of the NSPCC child maltreatment study (see Radford et al. 2011) questionnaire distributed which included a self-select tool to permit interviewees to be contacted for a follow up interview, therefore participants were self-selected or made up of	discuss the calculation of polyvictimisation within a national context, mindful that it is not an 'exact science, as there is currently no research consensus on the number of victimisations that makes someone a polyvictim' (p59). • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size –** n=60 interviews conducted with young adults aged 18–24.	but they don't really understand, they just don't understand' (female, emotional and physical abuse by mother up to the age of 17; p24). Young people (n=42) described poor family relationships as a factor, for example, 'one parent was perpetrating the abuse while the other facilitated the abuse or knew about it and did nothing to intervene' (p26). In some instances, siblings were being abused also, so were not always a source of support. Family stress also attributed to children not disclosing, partly because they had other factors that impacted negatively on the dynamic, i.e. disability, substance misuse and parental mental health problems. Conversely, some young people describe a positive family relationship (n=17), yet this impacted on the disclosure process because they did not want to upset the family dynamics. 'One young woman abused by her neighbourhood peer said that she knew her father would "go mental" and she was concerned that he would end up in prison for attacking her perpetrator in retaliation' (p26). 1.2. 'Perpetrator tactics': 21 young people described perpetrators actions which prevented them from telling, e.g. the 'perpetrator mask' which is where perpetrators manipulate others into believing the child is to blame. A female young person who was sexually abused by a neighbour and physical and emotional abuse by mother recounts: 'yeah, to begin with when I was fairly young, it was hard for my nan to see it like obviously. She just saw it like that I was a child playing up and that's why these things were happening. Because my mum was so good at twisting stuff, whoever it was I spoke to, it would be twisted around and I would just be like uh … (laughs), it's like that'.	not provide a representative of the general young adult population nor those who have experienced abuse. As stated in limitation 'although the young people in this study did volunteer to discuss their childhood experiences, the researchers do not believe this means these young people represent a sample that is predisposed to disclose their abuse. If this were the case, the latency period for disclosure would be much shorter' (p60). Overall assessment of external validity ++ Very comprehensive research design, with clear demographic and characteristics of participants (n=60). Overall validity score ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
convenience sample. The interview was two hours long and 'focussed on information that was not collected in the questionnaire in order to avoid duplication' (p58). **Country:** United Kingdom. **Source of funding:** Voluntary/Charity - NSPCC funded the project.		1.3. 'Developmental barriers': Some young people account that they didn't know what was happening to them or have the ability or vocabulary to ask for help. This theme was particularly common in those who experienced sexual abuse. One female who was abused by father from age 4 to 14 reported that 'it was just the norm for me sort of so I didn't think anything was sort of wrong with what was going on until I gotten older' (p27). Others considered physical abuse as part of parental discipline. 1.4. 'Emotional barriers and anxieties': A common theme amongst 12 young people who were sexually abused was feeling shame, guilt or embarrassment to tell others about the abuse. Mental health difficulties in the young people prevented them from reaching out for help (n=3). Anxieties about the reaction of others stopped 14 young people seeking help. One young person describes seeking help for depression and was cutting herself and trying to kill herself but was told 'go home and grow up' (female, physical and emotional abuse by mother; witnessing domestic violence; and being bullied at school; p29). 1.5. 'No one listened and no one asked': Some young people (n=4) account trying to ask for help but were not acknowledged. These young people describe poor relations with their primary caregiver. Compounded with 7 young people who did not disclose because they were not asked. One young person described '… no one asked me if I needed help and I think, looking back it was, like, I don't know, kind of the indicators you get if someone's being abused were there' (female, sexually abused by older cousin from age 8–12, p30).	A high quality study that meets the research aims and objectives and deals well with ethical concerns.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		1.6. 'Confidentiality': Seen as important to a young person disclosing, only a few young people had concerns about confidentiality and negative experiences when they confided in a professional (social worker and counsellor) who then relayed back to parent. Four young people describe not knowing that there were confidential services, i.e. ChildLine and if they had known, they would have used these services. 2. What promoted disclosure: With reference to young people's experience with professionals, some young people (n=12) stated that they had disclosed because someone had asked or noticed a behavioural change. One young person recalls being told by a social worker in hospital that if they told the truth about their bruises, they would be kept safe. Similarly, a young person lost a lot of weight and told their head of year that 'things weren't good at home' (female, sexual abuse by father from age unknown until 15; p31). 3. Informal disclosure through (3.1) friends and (3.2) mother: 3.1. Friends are described as being a 'critical support for many of the young people (n=38) in this study' (p35). Most sought help or support in seeking help, and emotional support. The study explores in-depth the way in which friends assist in disclosing the abuse, for example one young person states that she '… got forced to tell her [the teacher], 'cause one of my friends knew stuff was going on, so, got her in and we talked about it, erm, and erm, she told my head teacher' (p37). The findings suggest raising awareness amongst all young people of the dynamics and	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		consequences of abuse, as well as promoting messages in how to respond and ask for help. 3.2. According to findings, mothers were the recipients of 32 initial or linked disclosures. Yet only 5 young people report that their mothers took formal protective action by reporting the perpetrator to the police. The majority of young people experienced being ignored or 'attention seeking'. 4. Formal disclosure, which shines a spotlight on teachers, criminal justice system and social services: 4.1 Education: 18 young people made a total of 23 disclosures to teachers, which are characterised both positively and negatively. Positive experiences (n=6) are where education staff automatically believed the young person's disclosure and reported the abuse to the appropriate channel, i.e. head teacher or police. The experience is described as being straightforward however 'the aftermath included involvement by police and social workers, some of whom invited the young person's parents into the room during the interviews' (p40). Negative experiences of disclosing to teachers were characterised by the teachers' failure to inform the young person of how the disclosure would be handled. Named most unhelpful was the teacher going straight to the parents, some of whom were the perpetrators of abuse. One young woman who was living with domestic violence between her mother and mothers' partner said that the disclosure made things worse for her: 'I told one [teacher]. I wrote her a letter. I wanted someone to do something. She told my mum. I remember mum picking me up	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		from school and she dragged me out and shouted at me' (p41). 4.2 The criminal justice system: Over half of the young people (n=37) report involvement of the police. Findings suggest that generally young people found the process of disclosure 'uneasy and unpleasant'. Four key areas were highlighted to communication skills needed to handle information sensitively: being developmentally sensitive; being direct; allowing children to go at their own place; and keeping children and young people informed. - Developmentally sensitive: This is about treating young people with respect and being age aware. One young person commented: '… he was really good. Um, he talked to me like a young person, not like a stupid little kid' (female, sexually abused by step-father from age 7–16; p42). Conversely, not making assumptions that young people have the vocabulary to disclose. - Being direct: With relation to sexual abuse, one young person explains that she had two different police officers, one appeared to be embarrassed for example, 'and did he touch you down there?' whereas the other police officer who took the full interview, was direct asking proper questions. The author states that there are clear implications for training all police officers, not only those with special roles interviewing children. - Pace of interview: Some young people describe 'feeling out of control' as supported in one extract by a young person: 'It was, like, I was just waiting for somebody to, like, wake me up and I was going to go back into what my life was normally like. Um, the police interview on the Friday went on for hours', I think it lasted about four hours, um. Again the police, the two police officers were, they were really good, um,	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		and didn't like rush me, they didn't, like make me talk about things I really didn't, they just went at my pace' (female sexual abuse by step=father from 7–16; p43). - Keeping young people informed: 3 young people account that they wish they had been kept up-to-date as to what would happen after their disclosure, i.e. would the perpetrator be arrested and what happens next. 4.3 Social Services: 34 young people report that social services had been involved with their families in childhood prior to disclosure. These are relevant because they present a 'missed opportunity 'for professionals to react. 20 of the 34 reported that social services became involved after they disclosed the abuse to the police. Some young people (n=7) reported that they held negative assumptions and anxiety about the role of social services. As one young person recalls her family friend contacting social services: I think just 'cause I was scared 'cause like so many people try so hard like to keep it quiet then you just get scared by what will happen if you don't. And I didn't like, I didn't want bad things to happen if she had done (called them) and it kind of panicked me more what would happen, than like the situation that it was at the time' (p46). Twelve young people reported that social workers were working with their family, but some professionals did not engage with them (n=4); ask the 'right question' (n=5); and acted insensitively when they did visit (n=2). Findings suggest the importance of talking to children and young people on their own.	

2. Beckett H, Brodie I, Factor F et al. (2013). 'It's wrong … but you get used to it' - A qualitative study of gang-associated sexual violence towards, and exploitation of, young people in England. London: Office of the Children's Commissioner for England

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: 'The research aimed to consider: the scale and nature of gang-associated sexual violence and exploitation in six areas of England; the main pathways into gang-related sexual violence and exploitation for young people living in these neighbourhoods; and potential models for an effective multi-agency response to the issue' (p6). **Methodology:** Qualitative study. The research team adopted a qualitative approach: - Individual interviews with 150 young people - 11 focus groups with 76 professionals - 8 single sex focus groups with 38 young people. In relation to recognition, most relevant data was provided by research with children and young peo-	**Participants:** Children and young people. Individual interviews - Young people aged 13–28 (n=150). Focus groups - Young people (n=38). Professionals/practitioners - 11 focus groups were conducted with 76 professionals across 6 research sites. Representation from fields of social care, education, health, policing and the justice system, specifically working within the gangs and sexual exploitation/sexual violence. **Sample characteristics:** • Age - Interviews - Participants ranged from 13 to 28: Under the age of 18 (49%); 18–20 (28%); 21–25 (21%); and 25-28 (2%). Focus groups - Not reported. Professionals - Not reported. • Sex - Interviews - 52% were male, with 48% female. Focus groups - Not reported. Professionals - Not reported. • Ethnicity - Interviews - The self-reported ethnicity of interviewees: 32% Black/Black British; 28% White; 21% Dual heritage; and 18% Asian/Asian British. Focus groups - Not reported. Professionals – Not reported.	**Narrative findings:** Factors that hinder recognition are explored: 1) The acceptance of sexual violence and exploitation as a 'normal' part of life in their social milieu and a resignation to this; and 2) Low levels of reporting and seeking help from professionals (p43). 1) The acceptance of sexual violence and exploitation as a 'normal' part of life in their social milieu and a resignation to this: - Young people stated they felt resigned to the experiences within a gang environment. One young woman's response to findings demonstrate this, 'I'm used to it … it's normal … Welcome to our generation' (p43). - Professionals commented upon a resignation, as noted in one focus group, 'A lot of schools and charities I'm speaking to are like "how do we break this, how do we change this whole culture of how young people perceive their relationships?"' (p43). - Evidence indicates that young people (particularly woman) and some professionals feel there is little to be done to change the normalisation of sexual violence and exploitation within the gang environment. - The response from one young woman to end the sexual violence were limited to '(i) move out the area, (ii) get involved in another gang for protection or (iii) wait it out' (p43). This indicates the challenges to professionals seeking to prevent and respond to victimisation, if young people who do not see seeking external support as an option.	**Overall assessment of external validity:** ++ The study meets all criteria and has dealt effectively with ethical considerations. In addition, the interview guide was co-produced with the YPAG to make questions age appropriate. **Overall assessment of internal validity:** ++ Very comprehensive, effective study with detailed findings that are summarised into recommendations. The methodology has dealt appropriately with the ethics and risk associated with the subject matter, and the research team have ensured the voice of the child is at the heart of the report. **Overall score:**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
ple. There is a comprehensive methodology section (pp12–15) that addresses approach to interviewing; the research was co-produced with young people from the Young People's Advisory Group; information on gaining consent; and thorough steps taken to provide confidentiality and anonymity of participants. The study took place between 2011 and 2013. **Country:** UK. England. 'To maintain confidentiality and protect participants, the identity of the research sites is not being revealed' (p6). **Source of funding:** Government - Inquiry of the Office of the Children's Commissioner into child sexual exploitation in gangs and groups. Led by the University of Bedfordshire.	• Religion/belief - Interview - Not reported. Focus groups - Not reported. Professionals - Not reported. • Disability - Interview - Not reported. Focus groups - Not reported. Professionals - Not reported. • Long term health condition - Interview - Not reported. Focus groups - Not reported. Professionals - Not reported. • Sexual orientation - Interview - Not reported. Focus groups - Not reported. Professionals - Not reported. • Socioeconomic position - Interviews - 'Most participants reported that they were in some form of education (45%), training (20%) or employment (18%), with only one in eight identifying as Not in Education, Employment or Training (NEET). Focus groups - Not reported. Professionals - Not reported. • Type of abuse - 87% (n=131) had direct, often multiple connections with gangs. Of the 131 participants, 59% were/had been directly involved in a gang (M=70% v. F=47%); 32% had been gang-associated (M=25% v. F=39%); 35% had friends/and or family involved; 23% were having/had previously had a 'romantic relationship; with a gang-involved person (all female bar one); 57% had personal experiences of sex and/or relationships in gangs. The remaining 13%	2) Low levels of reporting and seeking help from professionals: - 1 in 12 young people felt that they would report, or talk about, experiences of sexual violence or exploitation. If they did disclose, rarely was it indicated that it would be a formal disclosure – 2/3rds indicate that they would make an informal disclosure to a peer. 2.1. Judgement by others - Young people stated that they might be reluctant to disclose experience of sexual victimisation for fear they might not be believed/judged. - One professional commented upon the barriers of reporting when working with victims (commonly young women), who cite reasons for young people not disclosing, 'because I'm going to be called a slag' or 'I'll lose all my friends, all the girls will find out and they won't want to be my friend anymore' (p.44). - Self-blame acted as a deterrent for young people not reporting incidents, for example, one 21 year old young woman in an interview stated: 'I think the main problem with like rape. For so long it's been portrayed as like in a way the woman's fault, because you were dressing provocatively. Oh you were drunk... and I think if a woman gets really drunk and that ends up happening, I think a lot of girls kind of blame themselves, like oh I was really drunk, blah, blah, blah, but they don't see what's actually happened, and they don't see that the guy is in the wrong. They kind of blame themselves for it' (p44). 2.2. Fear of retaliation - Some young people commented on gang retribution if they were seen as a 'snitch' or 'grass'. One 17 year old young woman stated 'let me give you an example	++ An excellent, thorough empirical study which meets its research aim and details implications for practice and policy on a local and national level.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	(n=19) participants grew up in gang-affected neighbourhoods. Focus groups - Not reported. Professionals - Not reported. • Looked after or adopted status - 38% of participants reported current or previous involvement with children services, although it is not clear what support this was. Focus groups - Not reported. Professionals - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Interviews - Not reported. Focus groups - Not reported. Professionals - Not reported. **Sample size** Interviews - 150 participants. Focus groups - 8 single sex with 38 young people. Professionals - total of 11 focus groups held with 76 professionals.	of why people don't [go to the police]. Because if you go to the police station and say "this gang member raped me' that gang member might be found guilty and go to jail, but remember he's part of a gang". So all the ones in the gang, 500 people, 400 people, will come back to you, to your house. Could go to your family's house, you know. So you might as well keep it on the low and move on with your life innit ... If you go to the police, that's the wrong move. That's the worst thing a person could do ... It'll come back cos with gangsters they got to win innit. They never give up' (p44). 2.3 Lack of faith in services ability to protect them - Compounded with the fear of retribution, young people commented on the inability of statutory services to protect them. This was generally due to the perception of the police, an absence of convictions and the need for long term protection. One 16 year old young man said, 'we don't believe that police are there to help us ... When I've seen my dad get arrested the police was hard on him, grips him up and push his hands behind his back and that, and it's not like he's resisting, they had him on the floor outside. There's no need for that. I've seen police taser people and I've been hit on my leg with a kosh, it's not necessary' (p.4).	

3. Burgess C, Daniel B, Scott J et al. (2012). Child neglect in 2011: an annual review by Action for Children in partnership with the University of Stirling. Watford: Action for Children

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The annual review process by	**Participants**	**Narrative findings**	**Overall assessment of internal validity**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Action for Children in partnership with the University of Stirling seeks to 'gauge the current situation with regard to neglect and monitor the effects of changes in national and local policy' (p5). The project team collated evidence through a variety of methods. **Methodology:** Mixed methods. The qualitative aspect of the annual review was conducting 12 focus groups across 6 local authorities with a total of 114 participants across a range of agencies: children's services; housing; health service staff; the police; education; and third sector agencies. The focus group data is most relevant to research question as provides in-depth data in response to 'how good are we at recognising children who are	• Professionals/practitioners – Qualitative: n=114 professionals; Quantitative: n=47 local authorities. **Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** Qualitative sample size: Research team held 12 focus groups with 6 (two areas were combined) local authorities: Local authority 1: n=12 Local authority 2: n=14 Local authority 3+4: n=47 Local authority 5: n=21 Local authority 6: n=20 Total participants = n=114 professionals.	How good are we at recognising children who are at risk of, or are experiencing neglect? 1. According to focus group participants, recognition has increased and there are more referrals. In part, it is accredited to fear of consequences of not referring, despite not having evidence. Referrals from primary schools have risen which could be 'due to better recognition, as one focus group respondent states: 'Although neglect is less clear-cut than other forms of abuse there is more awareness than there was an earlier identification. Staff from agencies, such as housing, know what to look for when they are going into homes for other reasons' (p9). 2. Focus group respondents highlighted 'the crucial role of nurseries, Sure Start children's Centres and health service staff, namely GPs, midwives and health visitors, in identifying also responding to young children at risk of neglect' (p10). A focus group respondent explained that this was due to parents having daily contact with these services, so will be best place to recognise signs of neglect. Are neglected children still not being identified? 1. Focus group professionals stated that there are still large numbers of children who are experiencing what may be called borderline neglect, thus falling beneath the threshold. As 1 focus group respondent states, 'There is good recognition now when neglect is more overt, but still not enough at the less obvious stage. Social workers need to be able to spend more time actually with families to really see what is going on in the home' (p11).	- The annual review has carried out 12 focus groups which include 114 representatives from different agencies, however the findings and conclusions are 'somewhat convincing' because there is difficulty in identifying or contextualising who said what. There is no consideration of limitations or theory underpinning focus groups. **Overall assessment of external validity** ++ Overall, study meets most of the quality criteria however the study is not co-produced. **Overall validity score** - The annual review meets the aim through the research design and mixed method data collection approach. The findings are representative of a

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
at risk of, or are experiencing, neglect?' Data from the other strands of the research are less relevant and have not been reported. **Country:** UK. **Source of funding:** Not reported.		2. Professionals distinguished the 'drip-drip' effect which is where there is not enough evidence to proceed: 'Some children are very good at masking what is going on and have well-developed defence strategies to cope with this. Apparent resilience can cover things up. Then a major incident occurs and it goes straight to child protection procedures. And children may hold a situation together through loyalty to their parent' (focus group respondent, p11). 3. Issues of capacity was stressed as an issue in that social care agencies are inundated with referrals and the child is not always being identified as quickly as they should be. Additionally, increased case load was seen as a reason social care staff have less time to spend with the family making children harder to identify: 'Higher caseloads will lead to less resources for each family' (p11).	large sample of professionals that work with children who are at the frontline for identifying and responding to child neglect. However, there is little information about consent of participants or what geographical region data is collected, so caution to generalise. Conclusions are difficult to see as reliable.

4. Children's Commissioner (2015) Protecting children from harm: A critical assessment of child sexual abuse in the family network in England and priorities for action. London: Office of the Children's Commissioner for England

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: Aim to assess the scale and nature of child sexual abuse in the family environment in England. We have extracted data in relation to Recognition (sections 14.1, 14.2 and 14.3)	**Participants** Adult survivors of child abuse. A survey of 756 survivors of child sexual abuse, all were over the age of 18. Professionals/practitioners - Site visits and focus groups in 6 sites, involving 32 agencies. Oral evidence hearings with 9 professionals from	**Narrative findings** Data have been extracted from sections 14.1, 14.2 and 14.3. 14.1 Recognition	**Overall assessment of internal validity** - **Overall assessment of external validity** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
and the Impact of intervention (section 15.3), as these relate most closely to our review questions. **Methodology:** Mixed methods. Study comprised: 1. A call for evidence to collect examples of good practice 2. A DfE dataset request for data on victims and perpetrators 3. Police force dataset request for data on victims and perpetrators 4. Site visits and focus groups in 6 sites, including consultation with 32 agencies and focus groups with 5 victim/survivor organisations. 5. Oral evidence hearings with 9 professionals from statutory bodies and 10 professionals from voluntary and community organisations 6. A survey of 756 survivors of child sexual abuse 7. Data request from 4 helplines 8. A rapid evidence assessment of	statutory bodies and 10 professionals from voluntary and community organisations. **Sample characteristics** • Age - Survivor survey: 18-24 n=50, 25-34 n=133, 35-44 n=214, 45-54 n=251, 55-64 n=88, 65+ n=20. Other evidence strands: Age of participants not reported. • Sex - Female n=483, Male n=51, Unknown n=215, Other n=5, Prefer not to say n=2. Other evidence strands: Sex of participants not reported. • Ethnicity - Not reported. • Religion/belief - Survivor survey: No religion n=283, Unknown n=215, Christian (all denominations) n=196, Other n=42, Jewish n=10, Buddhist n=7, Muslim n=2, Hindu presume n=0 (not shown on pie chart) Other evidence strands: Religion of participants not reported. • Disability - Survivor survey: No disability n=397, Disability n=106, Unknown n=211, Don't know n=41, Prefer not to say n=1. Not reported for other strands. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported.	The study reports that the focus groups with adult survivors and the call for evidence identified that recognition of intrafamilial sexual abuse may be delayed. Participants in site visits and oral evidence noted that disclosure may be 'involuntary' - when abuse is discovered or reported by a third party. Total 141 (26%) respondents to the survivors' survey reported that they did not recognise that they had been sexually abused until adulthood. Reasons given by adult survivors (unclear if via focus group or survey) for not recognising abuse included: - not being able to find the 'right words' (p60) to explain what had happened - sometimes this only occurred after seeing media coverage, after sex education or through conversations with friends as an adult - in some families 'sex is a taboo subject' (p60). The study notes that, even though not all children can recognise when they have been abused, adults may be able to identify that there is a problem based on their demeanour or behaviour. Of respondents to the survey, 293 were not asked by an adult about any of these factors, but 147 were. Survey respondents included behaviours such as becoming withdrawn (n=26), risk taking and aggression (n=20), alcohol/substance misuse (n=14) and running away from home (n=10). The study reports that the evidence it has considered suggests that victims and survivors are more likely to recognise what has been happening to them as abuse when they are in a 'safe space' (p.61) - characterised by presence of a trusted adult, and protection from the perpetrator. 14.2 Telling	**Overall validity score** - Little methodological information provided, particularly regarding survey distribution, response rates and representativeness of resulting sample. Limited consideration of ethical issues in reporting.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
research evidence on intra-familial sexual abuse 9. There is on-going research with children and young people (assume this is not reported here) The data extracted here are drawn from strands 1,4, 5 and 6. **Country:** UK. **Source of funding:** Government - Office of the Children's Commissioner.	• Type of abuse - All survivors had experienced child sexual abuse within the family. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** Survivor survey: 756 Site visits: 32 agencies (unclear how many individuals) Focus groups: 5 focus groups victim/survivor organisations and 3 focus groups with survivors of child abuse (unclear how many Oral evidence hearings: 9 professionals from statutory bodies and 10 professionals from voluntary and community organisations Total sample size unclear.	The study notes that evidence from the site visits, call for evidence, focus groups and oral evidence sessions suggests that victims of sexual abuse may 'tell' through their behaviour rather than a verbal disclosure. In the survivors survey many respondents (n=325, 43%) did not try to tell anyone. 234 people (31%) reported that they had tried to tell someone: 47% told 1 person, 32% told under 5 people and 20% told 5 or more people. Some respondents had tried to tell a professional, but disclosures were not always handled appropriately (although note that, due to age of some survivors, these experiences may go back 30 years or more). Total 44 respondents reported that they had tried to communicate through changing their behaviour including drawing pictures (n=10), playing with dolls in a particular way (n=4). When asked which person they had told, the responses were as follows: Mother (n=102, not possible to calculate percentage as base not provided); Friend/peer (n=85), Teacher (n=51), Father (n=32), Social worker (n=28), Sister (n=24), Police (n=23), Friend's parent (n=21), GP/doctor/nurse (n=21), Grandparent (n=17), Aunt (n=15), Brother (n=15), Cousin (n=12), Religious leader (n=9), Uncle (n=7), Other (n=42). 'Other' includes helplines, therapists/counsellors, partners/spouses. The study reports that, of the 220 people who answered the question 'did the abuse stop as a result of telling?'; 130 (59%) said it had stayed the same or got worse whereas 72 (33%) said it had stopped completely or temporarily. 14.3 Barriers to telling	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		The study reports that, across all forms of evidence gathered, the following themes were highlighted as barriers to telling: - Self-blame, in which victims think they are responsible for causing the sexual abuse - Guilt and fear of consequences, including loyalty to family members - Fear of the perpetrator - Fear of being judged - A lack of opportunities to tell someone - Distrust of professionals (p. 66). One respondent to the survey wrote: 'I didn't realise it was abuse until he'd made me do too much, by then I was so ashamed I didn't have the words or the confidence in myself ... I thought I would be in trouble and that I would hurt my family. I was just a little girl' (female survivor aged 35-44). 14.3.2 Telling professionals The study reports that the site visits and call for evidence identified that there may be additional barriers for children and young people to report abuse to professionals, including the fact that professional environments may be intimidating, and not all professionals have the right skills to talk to young people about these issues. There may also be language barriers including: - young people not having the right vocabulary or language skills - young people may have a learning disability or communication impairment - ethnic minority or asylum seeking children and young people may not know the right terminology to describe their experiences and seek help.	

5. Cossar J, Brandon M, Bailey S et al. (2013) 'It takes a lot to build trust' - Recognition and Telling: Developing earlier routes to help for children and young people. London: Office of Children's Commissioner

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: Research aims: 'To examine young people's perceptions of abuse and neglect, and to explore their experiences of telling and getting help from both informal and formal sources. - To use this knowledge to make suggestions for practice that would improve access to support' (pi). **Methodology:** Mixed Method. Mixed methods study comprising: - a systematic literature review about recognition and disclosure of abuse by young people, and their views of service (findings not reported here as review did not match our evidence criteria) - content analysis of an online peer support website for young people to post and respond to problems	**Participants** Children and young people. Content analysis of internet forum - analysis of 261 threads (corresponding to 261 individuals) Interview study - 30 young people aged between 11 and 20. Young people were deemed to be vulnerable, but were not currently involved with child protection services. Some had been involved with children's social care in the past. Focus groups - 1 focus group with 10 children aged 10 to 11 and one focus group with 10 young people aged 16-17. Caregivers and families - Via focus group study: Family focus group with 8 parents/caregivers. Adult survivors of child abuse - Some of the young people (number not specified) participating in the interview study were aged between 18 and 20. Professionals/practitioners - Via focus group study: One focus group with 8 community services representatives (1 deputy head teacher, 3 domestic abuse workers, 2 youth workers, 1 children's worker, 1 youth offending team worker); 1 focus group with 4 integrated services representatives (2 specialist health visi-	**Narrative findings** This an extensive report, with detailed findings presented in each chapter. The reported findings are therefore based on the executive summary, with an additional detail added in from each chapter as relevant. Findings from the literature review and two focus groups with young people (who had no experience of maltreatment or of using specialist services) have not been included here, as they did not meet our evidence criteria. 1. Recognition It is important to note that the term 'recognition' is used in this study to mean young people themselves recognising and acknowledging that they are experiencing abuse or neglect, rather than recognition by professionals. 1.1 Analysis of internet forum threads The study found that 23 'problem types' emerged from young people's descriptions. These were: '1) Mild sexual - sexual comments, partner initiating sex without asking; 2) Medium sexual - child being touched or made to touch perpetrator (genitals)/perpetrator watching child in bath etc. (age inappropriate); 3) Rape/serious sexual assault/buggery; 4) Media based – grooming via net, posting pictures or videos on web/social network, forcing child to watch pornography; 5) Sexual abuse – not specified; 6) Physical abuse – unspecified; 7) Severe physical – strangling/beating/causing injury; 8) Chemical physical –	**Overall assessment of internal validity** ++ **Overall assessment of external validity** ++ **Overall validity score** ++ Overall very good quality study with use of peer researchers, and good consideration of ethical issues. Good reporting of methods of analysis and clear findings.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
about abuse and neglect (261 threads) - interviews with 30 vulnerable young people, aged 11-20 who were 'currently deemed to be at risk, but who were not currently involved with child protection services' (p6) - 6 focus groups with children (data not reported here as do not meet our population criteria), young people (data not reported here as do not meet our population criteria), parents and practitioners. Parents were recruited from a parenting support group at a local voluntary organisation. **Country:** UK. **Source of funding:** Office of the Children's Commissioner.	tors, 2 early help practitioners); 1 focus group with 4 children in need representatives (2 social workers, 2 family support workers). **Sample characteristics** • Age - Content analysis of internet forum - Information not available Interview study - Age range 11-20 years, mean age 16 years. Focus groups - Children's workshop - 10-11 (Year 6); young people's workshop - 16-17 (Years 11 and 12). • Sex - Content analysis of internet forum - posters identifying selves as female n=162 (62%), posters identifying selves as male n=25 (10%), posters not stating their gender n=74 (28%). Interview study - female n=17; male n=13. Focus groups - Not reported. Focus groups - Not reported. • Ethnicity - Content analysis of internet forum - not reported Interview study - White British n=18, Black African/Caribbean/Black British n=9, mixed/multiple ethnicity n=3. Focus groups - Not reported. • Religion/belief - Content analysis of internet forum - not reported. Interview study - not reported. Focus groups - not reported. • Disability - Content analysis of internet forum - not reported. Interview study - not reported. Focus groups - not reported.	poisoning, drugging, forcing alcohol; 9) Physical abuse unspecified; 10) Inappropriate responsibilities (e.g. for household tasks, for care of a sibling; 11) Not enough food; 12) Parental lack of interest – limited engagement with child, e.g. prefers computer to child, no interest in child's achievements or worries; 13) Neglect – unspecified; 14) Verbal aggression – including shouting, insults, could also come by text; 15)Witnessing violence/domestic violence; 16) Child thrown out of house/locked out; 17) Criticism/blame; 18) Threats made to child/physical intimidation; 19) Sibling preferential – child thinks sibling is preferred to them, examples are 'mild' may be ordinary sibling rivalry, e.g. my brother has better trainers than me; 20) Carer damages on purpose/sells young person's possessions; 22) Parent unpredictable towards child; 23) Emotional abuse – unspecified'. The study reports that there were a number of factors that could prevent a young person from realising that their experiences were abusive or neglectful, including: a) The young person feeling that they deserved it. One young person posted: 'I believe every word said by my mum that I'm no good, that I'm useless, that I've done everything wrong' (p37). b) A difficulty in acknowledging that a parent could be abusive. For example one young person posted that they felt 'traitorous' (p37) in thinking that their treatment by the parent was abusive. c) A parent's unpredictability when abuse was episodic, and relationship was sometimes good. One young person posted 'Other friends of mine envy me for having 'such a great mother'. Usually until she's	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Long term health condition - Content analysis of internet forum - not reported. Interview study - not reported. Focus groups - not reported. • Sexual orientation - Content analysis of internet forum - not reported. Interview study - not reported. Focus groups - not reported. • Socioeconomic position - Content analysis of internet forum - not reported. Interview study - not reported. Focus groups - not reported. • Type of abuse - Content analysis of internet forum - physical abuse 63 threads (24%); sexual abuse 125 threads (48%); emotional abuse 50 threads (19%); neglect 23 threads (9%). Interview study - young people were asked to complete a questionnaire of problems they had experienced in relation to abuse and neglect. Other children calling you names, saying mean things to you n=17; other children hit you, jumped you or attacked you n=11; parent or carer ever beaten, kicked or physically hurt you or siblings n=10; parent or carer ever beaten or kicked or hit you with an object like a stick or wooden spoon n=8; parent or carer calling you names, saying mean things n=8; another young person tried to force you to do sexual things you didn't want to do n=8; boyfriend or girlfriend slapped or hit you n=7;	angry. Most of the time she denies drinking … she doesn't even remember a large portion of the things she does' (p37). d) Confusion about the boundaries between discipline and physical abuse. e) Confusion about boundaries relating to touching with family members. One young person posted about a male member of the family who made her feel uncomfortable because he wanted her to sit on his lap: 'He might be just showing affection and I don't want to make a big deal out of it if I've got it all wrong, but it does make me feel really uncomfortable' (p39). 1.2 Interview study The study reports that the interviews identified a range of responses by young people including lack of recognition, partial recognition and clear recognition. a) Lack of recognition This included being 'actively dismissive' (p61) where young people did not agree with 'adult' definitions of abuse, e.g. in the case of sexual relationships between peers or in being a young carer. It also included a lack of awareness, often linked to age. Some young people said that 'closing off' from the abuse was central to their way of coping. One young person said 'I was just like "I don't need this, I don't need nobody, I will be alright"' (p62). b) Partial recognition Some young people reported that they had partially recognised the abuse, e.g. by feeling uncomfortable about a situation, but had not been able to identify	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	anyone ever stolen anything from your house n=7; anyone used force to take something away from you that you were carrying or wearing n=6; parent or carer shaken you very hard or shoved you in to a wall or piece of furniture, n=6; after break up parents taken you, kept you, or hidden you to stop you being with the other parent n=6; grown up ever forced you to have sex when you didn't want to n=4; grown up ever touched your private parts when they shouldn't have or made you touch their private parts n=3; ever attacked because of skin colour, religion, origin, physical problem disability or because someone said you were gay n=3; neglect n=2; going to school regularly in clothes that were torn, dirty or did not fit n=1. Focus groups - not reported. • Looked after or adopted status - Content analysis of internet forum - not reported. Interview study - not reported. Focus groups - not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Content analysis of internet forum - not reported. Interview study - not reported. Focus groups - not reported.	this as abuse. Young people talked about gradual realisation of the situation. However, they often did not know how to change things, even once they had recognised that they were experiencing abuse or neglect. c) Clear recognition There were some examples where young people knew at the time that they were experiencing abuse – this was more likely when the abuse happened in adolescence. 2. Telling 2.1 Analysis of internet form threads The majority of posts on the internet site did not state any barriers to disclosing abuse (58% of threads). The study found that young people posting on the internet forum reported the following barriers to telling. These were (in order of frequency): - Emotional barriers such as 'shame, embarrassment, not being able to face telling, finding it hard to find/say the words' (pv) (21% of threads). - Worries about their family knowing, loyalty to their families and impact on their family member (11% of threads). - Concerns that their situation was not sufficiently serious to tell anyone about (7% of threads). - Being threatened by their abuser (6% of threads). - Fear of not being believed (5% of threads). - Fear of negative consequences (5% of threads). - Fear that their information would not be kept confidential (3% of threads). - Protecting a non family abuser (1% of threads).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		- Fear of stigma or ridicule (1% of threads). - Practical constraints (1% of threads). Of the 261 threads analysed, there were 109 in which it was clear that the young person had told someone about the abuse. Where abuse had been disclosed this was to a professional in 41% of cases, to a friend in 31% of cases and to a family member in 28% of cases. 2.2 Interview study Young people in the interview study talked about who they would tell about abuse. The results suggested that the choice of who to tell was related to the type of help that was wanted. The results for each group are shown below, numbers in brackets relate, respectively, to: 'Stop abuse or problem', 'information and advice', 'emotional support', 'practical strategies to minimise harm' and 'medical help'. Professional sources of help: Doctor (Stop abuse or problem: 1, Information and advice: 3, Emotional support: 2, Practical strategies to minimise harm: 1, medical help: 14); teacher (15,10, 7,6,1); social worker (21, 4, 6, 1, 1); teaching assistant (1, 2, 3, 0, 0); camhs (1, 3, 13, 5, 2); youth worker (10, 7, 10, 4, 0); youth worker (10, 7, 10, 4, 0); school nurse (0, 6, 6, 0, 8); church or religious worker (1, 0, 0, 1, 1); police officer (23, 1, 1, 0, 1); sports coach (3, 1, 1, 1, 0); helpline or internet (4, 9, 3, 0, 0); school (unspecified (0, 0, 0, 0, 0) counsellor (0, 2, 9, 1, 0); solicitor (0, 1, 0, 0, 0).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Informal sources of help: friends (peers) (3, 8, 18, 2, 0); friends (adult) (6, 0, 5, 2, 0); boyfriend/girlfriend (2, 2, 7, 1, 0); parents (12, 2, 6, 0, 0); siblings (2, 1, 2, 0, 0); extended family (5, 1, 5, 1, 0); family member (unspecified) (8, 3, 3, 0, 0). Key findings about who was perceived to be able to provide each type of were as follows: - Stopping the abuse – young people were more likely to cite the police (23 mentions), social workers (21 mentions) and teachers (15 mentions). - Information and advice – young people were most likely to mention teachers (10 mentions), helplines (9 mentions), youth workers (7 mentions) or friends (8 mentions). - Emotional support – friends were the most cited in relation to this role (18 mentions). The professionals most likely to be mentioned were CAMHS (13 mentions) or youth workers (10 mentions). - Practical strategies to minimise harm – Young people most often mentioned teachers (6 mentions), CAMHS workers (5 mentions), youth workers (4 mentions). - Medical help – doctors were most frequently mentioned (14 mentions) and school nurses (8 mentions) were perceived as having a wider role for emotional support as well as medical help. Four themes in relation to telling emerged from the interviews. These were: 1) Being 'hidden' by actively avoiding telling or passively not telling – this was related to wanting to be loyal to family or fears of being removed from home. One young person said: '… Mum getting arrested …	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		who known maybe my mum could be classed as not well enough to look after me and then I could get put in Social Services and then in just goes on and on and on … I would rather just stay with me mum no matter what the consequences' (p67). Some young people feared that telling would make the abuse worse. One young person said: 'I really, really wanted to open my mouth and just tell them everything and just get out of there, but I knew that I would have to go back to my mum on that night and then she would have definitely heard about that and I would have had a bad experience and I had a massive fear of that, so I dared not mention anything' (p67). The study found that some young people remained hidden because they were not noticed, e.g. 1 young person said 'I weren't surprised [that no one asked] because of the school I went to … they focused more on your school uniform than who you are' (p68). 2) A trigger point – 'sign and symptoms' which signal the problem and may alert others to the child. The study reports that in many of the interviews young people described an incident which led to the involvement of a 'helping person' (p. 68). This included externalising behaviour such as physically or verbally harming others. For example, 1 young person described violently assaulting her boyfriend: 'I lashed out on him, a year's worth of anger came out on him and he ended up in hospital and I ended up getting in trouble' (p69). Incidents called also include internalising behaviours such as self-harm or suicide attempts.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		3) Prompted telling Young people in the study talked about the importance of a 'sensitive and effective response' (p69) to the kinds of behaviours which could serve as trigger points which would encourage them to tell. Telling depended on the relationship with a helping person including 'trust, duration of relationship and closeness' (p69). One example given was of a young person who'd been sexually assaulted by a stranger. The young person had not told anyone about the incident. When she became very upset at a team around the child meeting, her teacher responded to her distress and showed persistence in persuading her to disclose the abuse. The young person said: 'I was upset that past week and she asked me how I had been since the one before and I said I had been fine and Miss said, "Well that is not completely true because the last week has not been so good as it could have been" and then I just started crying and she asked what was wrong and I said that I couldn't tell her. Then everyone else went and she stayed and she said. "You can always tell me anything, because you normally do, so whenever you are ready just go for it.'" (p70). 4) Purposeful telling The study describes some examples in which the young person had recognised the abuse and actively sought help – this was less usual than examples of prompted telling.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		The study reports that active telling depended on having a trusting relationship with the person who was approached. One young person said 'I was so close to her [teacher] and I spoke to her about everything … she was like my best friend, she was the only person that I would go to' (p70). Some young people described seeking help immediately, others were not able to tell until 'after a build-up of emotional tension, a critical point was reached which precipitated telling' (p71). Some young people first approached an 'intermediary' (p71) when they wanted help, for example telling their parents about bullying in the hope that their parents would tell teachers at school. Other 'indirect' methods of telling included writing letters or text messages. For example, one young person described texting her father to tell him she had been sexually abuse: 'When I did tell him I didn't tell him to his face, I wrote him a text because I couldn't look him in the eye!' (p71). 2.3 Focus groups Practitioners in the focus groups also emphasised the importance of professionals noticing children's distress through their behaviour, rather than placing responsibility on children to tell about the abuse. Professionals reported that if children are asked questions in a sensitive way then they may disclose abuse. This is likely to depend on there being a professional who young people trusted.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Participants in both the Family Focus Group and integrated services thought that it may be particularly difficult for children to speak about neglect because they would 'consider it normal because that's their life' (p95). 3. Help 3.1 Analysis of internet forum threads The study reports that of the 109 young people on the internet forum who had told someone about the abuse, 48 (44%) mentioned at least one positive outcome from telling, and 50 (46%) mentioned at least one negative outcome. Twenty young people mentioned both positive and negative outcomes. The positive consequences mentioned were: - A positive response from agencies. - A positive response from friends and families. Negative consequences included: - Not being believed. - Friends and family were not supported. - Authorities were not supported. - Abuser's behaviour not punished – e.g. 1 young person said 'I'm scared he's going to kill me because I've broken my promise' (p48). - Stigma/labelling. - Breaking confidentiality – e.g. 1 young person said about a friend breaking their confidentiality 'In the end he told someone, at first I felt really upset but I know he only did it because he wants the best for me' (p49).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		The study also suggested that users of the online forum found it helpful. Those responding on the forum often gave emotional support, reassurance that the poster was not alone in their experiences and recommended telling. 3.2 Interview study 1) Self-help The study reports that many of the young people said that they had developed their own strategies for managing their difficulties including distracting themselves from problems, and finding ways to be self-reliant. 2) Help from others The study describes young people's experiences of services. They described the following qualities of people and helping relationships as important: - Trust – young people were more likely to trust professionals than people in their 'informal network' (p73) but 'trust was hard to establish and fragile' (p73). Young people thought that trust was promoted by the duration of relationships. For example, 1 young person said 'It does take a long time to get to know someone and you know you can trust them, I mean I was there for four, four of five years, so I had known her [teacher] really well so that's why I went to her, I should have done it in the beginning it might have all stopped' (p75). Trust was also thought to be fostered by not being 'judged' by professionals and being believed, sensitive treatment of confidential information. - Accessibility and availability – this could be in terms of location, for example 1 young person said:	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		'She [support worker] is like in the middle of the school so all you need to do is just go downstairs and on the right she is there in the office' (p79). Accessibility was also related to time, for example the numbers of sessions available with a professional. - Closeness/distance – young people talked about the relative benefits of people to whom they were close, compared to those who were not too closely involved with the situation and could be more objective. - Knowledge, expertise and effectiveness – the study found that young people valued professionals who they perceived as having good expertise. For example 1 young person said: 'Social workers obviously take care of children and they care about that they are safe and being looked after, so that's why it came to mind' (p81). 3.3 Focus groups 1) Parents' views of help The study reports that parents had a range of views about where they would seek help, but that many said that they would not want to go to children's social care and would prefer to seek help from schools, the police or their GP. One parent said of children's social care 'I don't think there is any compassion, they are not interested in how you are feeling, they are just interested in the children' (p97). Some parents reported getting better help from the voluntary agency from which they were currently accessing support. One parent said: 'Well they have groups like this where everyone discusses open you	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		know what is going on and I think we learn a lot about each other and you realise that you are not the only person that is there' (p98). 2) Practitioners' views of help The study reports that practitioners identified several barriers to help. These included: - The skills of school staff to speak to children and young people about sensitive issues. - Schools may not speak to parents when they make a referral to children's social care. - Parents persuading workers to close a case. - Younger children may be deemed to be less capable of making their own decisions about requiring help, if parental consent is not secured. Practitioners also mentioned systemic and organisational barriers such as variations in thresholds, lack of consistency and pressure on children's social care. Conclusion The study authors used the findings from the study to develop a 'conceptual framework for understanding recognition, telling and help from the point of view of the child' (pix).	

6. Coy M (2009) Moved around like bags of rubbish nobody wants: how multiple placement moves can make young women vulnerable to sexual exploitation. Child Abuse Review 18: 254–66

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim	**Participants** • Adult survivors of child abuse -	**Other/general** **• Narrative findings**	**Overall assessment of external validity**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Exploratory study that aims to identify 'how local authority care places young women at risk of sexual exploitation through prostitution by unpicking the "culture of care", and focuses on frequent placement moves and the role of discontinuity' (p254). **Methodology** • Cross-sectional study Ethnographic study. Autobiographical accounts from interviewees using an unstructured interview technique. **Country** • UK. London. **Source of funding** • Not reported.	n=14 women. **Sample characteristics** • Age - 17–33 at the time of interview. • Sex - All female. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Statutory guidance defines the young women as victims of abuse as they all began selling sex under the age of 18 (Department of Health 2000), and 'all retrospectively framed their entry into prostitution as abusive' (p258). • Looked after or adopted status - Reasons for entry into local authority care included: disruptive behaviour (4); sexual abuse (3); physical abuse (2); neglect (2); abandonment at birth (2); and police protection when found selling sex at 13 years old (1). • Unaccompanied asylum seeking, refugee or trafficked children -	Findings relevant are: 1. Multiple placement moves -Young women echoed other findings to suggest that multiple placement moves negatively impacts on their care experience. One woman commented: 'Being moved around, that was the biggest issue with it [care]. As soon as you start to get settled somewhere that was it you was off, and moved somewhere else' (Jo, 30, p259). -All interviewees commented on a high turnover of placements, for 2 women, they could count 35 changes of foster care placement over periods of 7 and 4 years. Women described distressful feelings towards these foster care placement moves. One young women describes moving 200 miles away from her family. -Frequent moves was noted to disrupt school placements and relationships formed. 2. Hearing young women's voices in placement planning. -This section details the experience of Becky, who was moved without her approval (from her nan's to her uncle's and then away to a foster paper she disliked), which consequently led her to abscond. She met an older man who introduced her to heroin and groomed her into selling sex. -The author suggests to involve 'young women in decision making about placements and reducing the frequency with which they are moved may not only help to decrease the risk of finding attachments with predatory older men, but also increase the likelihood that they will accept support from agencies' (p261). 3. Searching for ways to settle - Not relevant. 4. Vulnerability to sexual exploitation. A possible outcome of instability.	+ Sound ethical consideration, however study only in part relevant to a research question (potentially recognition – risks). **Overall assessment of credibility (internal validity)** + **Overall score** + Sound empirical study with thorough research method and justification. However, caution to generalise findings as represent 14 adult survivors experience of sexual exploitation in London. The potential for research bias, although described in the study, is not detailed in limitations.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Not reported. **Sample size** n=14.	-Findings suggest from the small sample of women's narratives ways the psychosocial impact of placement instability precipitates entry into prostitution. - Participants in the research report how changes in carer prevented them from forming bonds with adults associated with professional roles, and led them to developing relationships with older predatory men and friendships with those within the street prostitution community. -The author notes a women's powerlessness and willingness to appease, for example one young women (Lisa, 21), reports: 'Every place they sent me, I tried to be good, you know, behave good and that, not play up, but it still never worked out for me' (p262). The author suggests that becoming a prostitute gives women a sense of financial stability and a way of exercising personal power over their lives (p263). Stability and security as potentially preventative measures. -All women were asked what could have prevented them from entering prostitution, to which one young woman (Christiana, 21) comments: 'They can show some love or caring, instead of this "we're moving you there". They need to stop moving people round like a bag of rubbish' (p263).	

7. Daniel B, Taylor J, Scott J (2010) Recognition of neglect and early response: overview of a systematic review of the literature. Child and Family Social Work 15: 248–57

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The systematic review of literature examined an initial 20,480 items with inclusion filtered	**Participants** Children and young people. Inclusion criteria included children aged from birth to 19. Populations of each included study not given. None	**Narrative findings** The findings provide a narrative summary of 5 authors studies in relation to recognition:	**Overall assessment of internal validity:** -

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
to 63. The aim was to examine the 3 questions posed by authors: 1. What is known about the ways in which children and their families directly and indirectly signal their need for help? 2. To what extent are practitioners equipped to reorganise and respond to the indications that a child's needs are likely to be, or are being neglected, whatever the cause? 3. Does the evidence suggest that professional responses could be swifter? (p248). We have reported only the findings in relation to points 2 (recognition) and 3 (response), as the data reported in point 1 do not meet the evidence criteria for our related review question. **Methodology:** The systematic review was based on review guidelines (Centre for	of the studies reported here (Appleton 1996; Bryant and Milsom 2005; Lewin and Herron 2007; Paavilainen et al. 2002; Rose & Meezan, 1995, 1996) appeared to involve children and young people. Caregivers and families - Inclusion criteria included parents, or both parents and children. Populations of each included study not given. Two of the included studies appeared to involve mothers (Rose and Meezan 1995, 1996). Professionals/practitioners - Inclusion criteria included studies of professionals. Studies reported here involved the following professional groups: Appleton 1996 - health visitors Bryant and Milsom 2005 - school counsellors Lewin and Herron 2007 - health visitors Paavilainen et al. 2002 - staff in a children's hospital Rose & Meezan, 1995, 1996 - 'professionals' (not specified). **Sample characteristics** Sample characteristics • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not report-ed. • Sexual orientation - Not reported.	- Rose and Meezen (1995, 1996) found that mothers were generally shown to express greater concerns about maltreatment than professionals in the UK and USA. The study reports that: 'Such studies suggest that the general population is at least as well-equipped as professionals to recognize aspects of neglectful care, if not more so' (p252). - Paavilainen et al. (2002) study in Finland suggested that 2/3rds of 513 staff in a children's hospital could recognise maltreatment, despite associated difficulties. - A questionnaire conducted in the UK with 92 health visitors asked participants to rate the importance of 45 signs and symptoms of neglect (Lewin and Herron, 2007). The study found that 'there was considerable agreement about the five signs and symptoms that were rated as most serious and the findings suggest that health visitors are equipped to recognise the importance of the parenting and emotional aspects of neglect' (p252). - Appleton's (2006) study found that health visitors are equipped to identify a wide range of vulnerable children which might not be recognised using a formal criteria. - One study examined the role of school counsellors in identifying neglect. - The authors note that there is no body of evidence about the police's role in recognising neglect (p309).	**Overall assessment of external validity:** + Thorough systematic review with clear criteria and guideline. However, the studies included are international. With the relevant section about recognition, the focus is different to guideline question. **Overall validity rating:** - Extensive systematic search, however little information given about individual included studies, and method for synthesising study findings very unclear.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Reviews and Dissemination 2007) where 14 bibliographic databases were searched with an inclusion of national and international primary research studies published in English from 1995 - 2004. There were originally 20,480 possible papers, reduced to 1532 by removing duplicates/unrelated items. Titles and abstracts were then screened by the research team which reduced to 686 (inter-rater reliability of 95%), which reduced further to 112. The sample of 112 was then read in full using a data extraction form and research was included if deemed a good paper by a guidance and criteria set out for each type of study (Scharr 2006). Of the remaining items, 63 papers were read in detail and analysed (p249). We re-	• Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** Appleton 1996 - not reported Bryant and Milsom 2005 - not reported Lewin and Herron 2007 - n=92 Paavilainen et al. 2002 - n=513 Rose & Meezan, 1995, 1996 - not reported. **Systematic reviews - number of studies -** The number of studies included a final dataset of 63 (from 20 480). We report here on the findings of 6 of the included studies (see above).		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
port here on the findings of 6 of the included studies (Appleton 1996; Bryant and Milsom 2005; Lewin and Herron 2007; Paavilainen et al. 2002; Rose and Meezan 1995, 1996). **Country:** Range of countries. Studies reported here were from: Appleton 1996 - not stated Bryant and Milsom 2005 - not stated Lewin and Herron 2007 - UK Paavilainen et al. 2002 - Finland Rose & Meezan, 1995, 1996 - USA and UK Systematic review was carried out by research team in UK (Scotland). **Source of funding:** Not reported.			

8. Gilligan P, Akhtar S (2006) Cultural Barriers to the Disclosure of Child Sexual Abuse in Asian Communities: Listening to What Women Say. British Journal of Social Work 36: 1361–77

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To give voice to the views of ordinary Asian women and of front-line Asian women workers, in the context of discussions about child sexual abuse. **Methodology:** Qualitative study, 12 focus group discussions with Asian women. **Country:** UK, England (Bradford). **Source of funding:** Not reported.	**Participants** Caregivers and families - Asian women in communities. Professionals/practitioners - Front-line staff from 12 organisations. The majority were Asian women; 3 were men. **Sample characteristics** • Age - Women aged between 20 and 60 years. • Sex – Female. • Ethnicity - Discussions took place in Urdu, Punjabi, Bangla and English; 90% of participants were from Urdu/ Punjabi-speaking communities. The remainder were from Bangla-, Pushto- and Gujarati-speaking communities. • Religion/belief – 'All appear to have been Muslim, but participants were not asked for individual information'. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Sexual. • Looked after or adopted status - Not reported.	• **Narrative findings** 'Cultural imperatives arising from concepts such as izzat (honour/respect), haya (modesty) and sharam (shame/embarrassment) are, for many in Asian communities, crucial determinants of behaviour in response to incidents of child sexual abuse' (p1367). Izzat, haya and sharam or associated concepts and equivalents in English were discussed by almost all the groups. These concepts meant that, despite professional attempts at 'confidentiality', victims and non-abusing parents may feel that the disclosure of sexual abuse is a public event. One participant said: 'If you've got white social workers turning up at the door all the time ... it's really hard then to keep it within that family to deal with it because the word kind of gets out in a community and you have to start explaining what's going on' (p1368). Some women in Bradford suggested that the cultural, and more particularly the religious, imperatives of their communities provide important possible foundations for appropriate responses to child sexual abuse - reinforcing arguments in favour of giving religious issues greater prominence in qualifying and other training for professionals. The paper discussed results in terms of these key themes. 1. Outreach work. Almost all participants in the consultations welcomed the fact that outreach work on the issue was taking place. They offered many practical suggestions for written materials, e.g. including child sexual abuse information within parenting training provided by Sure Start projects, and social work	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity rating:** - Research aims unclear, and study appears to bring in other sources of information such as practitioner data. Little information on sampling and methods.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** Twelve group discussions involving a total of 130 Asian women were facilitated at community centres and similar places in Bradford. Five groups gave consent to at least some of their discussion being recorded. These recordings were transcribed and translated. Forty people attended a consultation event, from 12 organizations. The majority were Asian women. Three group discussions were conducted in English and another in Urdu. Three groups were 'women only'. The fourth included the 3 men attending the event.	students on placements undertaking associated projects. Many also emphasised the importance of targeting both children and men in future outreach work. They suggested that this be done in schools and through existing men's groups and mosques. Many judged that parents and carers would be encouraged to take the issue seriously if they were given information which raised their awareness of the emotional and behavioural issues that the child would be going through in sexual abuse and after disclosure, and of the impact on children of not feeling heard. Others stressed the need to highlight the view that 'good' parents protect their children and to underpin this message with religious and cultural injunctions. 2. Training and support for practitioners Practitioners involved in the consultation event were particularly keen that they should receive more training in how to deal with disclosures and that there should be a more readily accessible network for advice and consultation. They saw this as combining the experience of those involved in sexual abuse prevention with the expertise of others on their own cultures. They also recognised the importance of colleagues becoming familiar with and sensitive to particular aspects of Asian cultures. These included not only very widespread cultural imperatives such as izzat, but also issues that may be relevant in work with specific groups. All practitioners need to remain alert to the fact that they are at risk of misinterpreting the behaviour of people whose culture they are not familiar with, particularly if they approach such situations with an insufficient awareness of their own potential for ignorance. They also have particular implications for the	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		qualifying and post-qualifying training of social workers and emphasise the need to ensure that development of culturally competent practice is prioritized at all levels. 3. Flexibility in service provision Many participants emphasised the need for practitioners to have sufficient discretion and to develop sufficient confidence to allow them to respond flexibly to people making disclosures. This reflected a perceived need for families to feel they are in control of the process and of the pace of events. They emphasised the need for practitioners to talk with families about what would happen if they did disclose and about the processes involved. In relation to children, they noted the need not only to advise about what will happen, but also to build confidence that disclosing will actually be helpful. They emphasised the desirability of families having only 1 or 2 professionals to deal with and the need for service users to be reassured that interpreters will maintain confidentiality. Others pointed to the need for non-abusing carers to be able to access advice, without other family members knowing. They suggested that 'drop-in' sessions could be provided at agencies such as Sure Start, but again recognised that workers in such sessions would need confidence, discretion and adequate support to make judgements about whether they needed immediately to alert investigating agencies to incidents of sexual abuse or whether such a decision could be left in the control of service users. The consultation participants noted that children will be better served if service users are able to talk anonymously, before reaching a point where they feel comfortable in taking matters forward.	

9. Harper Z, Scott S (2005) Meeting the needs of sexually exploited young people in London. London: Barnardo's

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The aim of the study was to understand: 1. The nature and extent of sexual exploitation in London 2. The service needs of young people at risk of sexual exploitation 3. Gaps in existing service provision in London 4. Examples of promising practice which could be shared across London. The study also includes examples of facilitators and barriers to identification of, and response to, child sexual exploitation, which is what our data extraction has focused on. **Methodology:** Qualitative study. Qualitative interviews with young people and practitioners as part of a wider study which also included audit. Young people's data	**Participants:** Children and young people. Twelve young people aged between 13 and 19. Professionals/practitioners - Interviews with a range of practitioners including child protection co-ordinator (n=32), police (n=10), health service (n=10), education service (n=2), local authority looked-after children's service (n=2), residential home manager (n=1), youth offending team (n=3), secure unit manager (n=5), specialist sexual exploitation service (n=6), voluntary sector service with expertise in trafficking (n=6), homelessness/going missing service (n=4), drug and alcohol service (n=3), adult sex worker service (n=3), other voluntary sector service (n=3). **Sample characteristics:** • Age - Young people: Aged between 13 and 19. Practitioners: Not reported. • Sex - Young people: 11 women and 1 man. Practitioners: Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported.	**Narrative findings** 1. Police - Identifying and preventing sexual exploitation The study reports that many of the respondents to the study noted that, although the police carry out some proactive work to identify perpetrators of sexual exploitation, that there could be more active identification and prevention. Work to identify and prevent was carried out through vice units, work with young people who run away or go missing and partnership working with social services departments. 2. Social services and ACPCs (former terminology for LSCBs) 2.1 Identifying child sexual exploitation The study reports that social services departments largely identified sexual exploitation from amongst children already known to them. One respondent said: 'There may be other children, but we don't get referrals and don't do any outreach, nor are there other young people's services that young people would access and we would get referrals' (child protection co-ordinator, social services, p61). (However, note that this study is relatively old - 2005 - and practice may have moved on.) It was noted that it was relatively unusual for social services to receive referrals of sexual exploitation. Local authorities with access to specialist services noted that these services enabled identification as they facilitated disclosure by young people.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** + Study was conducted in 2005, which means the findings may be somewhat outdated as awareness of, and practice in relation to, CSE has changed considerably since that time. Relatively sparse reporting of interviews with children and young people.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
relates only to Response. **Country:** UK, England. **Source of funding:** Voluntary/Charity - Corporation of London's Bridge House Trust.	• Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Young people were recruited from services that worked with young people experiencing or at risk of child sexual exploitation. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** Young people n=12 Practitioners n=90	2.2 Barriers to identifying young people at risk The study notes that practitioners acknowledge that young people often do not disclose this type of abuse. Respondents also thought that many practitioners were not looking for signs of sexual exploitation. Respondents noted that 'challenging behaviour' can be an indicator of sexual exploitation. Some respondents noted that workers may be reluctant to identify sexual exploitation because they did not know how to respond once it was identified. 3. Health, education and other statutory services - Identifying and preventing sexual exploitation 3.1 Health The study notes that difficulties were mentioned regarding worries about breaching the confidentiality of sexual health services in order to highlight issues of CSE. It was thought this could be ameliorated by good partnership working and providing clear guidance on thresholds for child protection. 3.2 Education The study comments that the overall ethos and atmosphere of schools can affect whether they are supportive of identification/disclosure of CSE. Barriers to identification included: - Homophobia in schools, which could make it difficult for young people to discuss sexual experiences openly. - Varying levels of staff awareness of CSE. Disengagement from school was noted as a risk factor/indicator for CSE, but the study notes that this is	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		harder to monitor in areas with highly mobile populations. 3.3 Youth offending teams Those interviewed had not identified young people at risk of sexual exploitation. 4. Supporting young people who have arrived from abroad - identifying sexual exploitation The study notes that young people may not always disclose that they have been trafficked and/or sexually exploited: - Unless they have built up a trusting relationship with someone. - Because they may not realise that their experience is 'trafficking'. - Out of fear of their abusers and the UK authorities. Lack of awareness amongst services was also cited as a barrier to identification.	

10. Kazimirski A, Keogh P, Kumari V et al. (2009) Forced Marriage Prevalence and Service Response. London: Natcen

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The research had two aims: 1. To improve understanding of the prevalence of FM 2. To examine how services are currently responding to cases of FM. The study states that is had 'a particular focus on UK resident children and young people under 18 years of age' (p1.) We have extracted data only in relation to research question 2, which has content which relates to our review questions 6 (Recognition), 14 (Early help) and 20 (Response). **Methodology:** Qualitative study. In-depth interviews with 40 key stakeholders across four case study local authorities. **Country:** UK, England.	**Participants** Professionals/practitioners - 40 professionals across 4 local authorities, covering both statutory and voluntary agencies. respondents included: statutory sector respondents • police - detective inspectors, superintendents, sergeants • domestic violence (DV) - DV community safety unit (CSU) officers, DV outreach services, DV co-ordinators • child protection (CP) staff - directors of children's services, local safeguarding children's board (LSCB) co-ordinators, safeguarding children co-ordinators, CP advisors and co-ordinators • education - education welfare officers (EWOs), school counsellors, student services officers, personal advisors • local councillors • primary care trust (PCT) public health managers • housing services staff. voluntary sector respondents: • black/minority ethnic (BME) and DV - DV women's groups staff, refuge staff, counselling staff • victim support workers • law centre workers • youth/children's charity workers • religious leaders (p13). **Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported.	**Narrative findings** These findings have been extracted from Chapter 5 on 'Detection'. 5.1 Forced marriage referrals The study states that most young people were identified to schools, college, youth agencies and BME DV/FM voluntary sector either by self-referral or through a concerned friend. The study states that direct reporting to the police was less common. Statutory and voluntary sector DV agencies reported that forced marriage usually emerged as an issue after a period of contact with a victim, rather than at their point of contact. Similarly, forced marriage often emerged as an issue via health services dealing with eating disorders and self-harm. Being missing from education was also considered a useful indicator of risk of forced marriage. 5.2 Factors preventing detection The study reports the following factors as preventing detection of FM. 5.2.1 Varying perceptions of forced marriage prevalence There was a discrepancy across survey respondents regarding perceived prevalence of forced marriage. Those in the statutory sector tended to have identified relatively few cases, whereas voluntary and commu-	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Good relevance to question, but no consideration of ethical issues. **Overall validity rating:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Source of funding: Government - Department for Children, Schools and Families with support of Forced Marriage Unit.	• Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** Forty individuals interviewed across 4 case study local authorities.	nity sector organisations (VCOs) and domestic violence organisations estimated the prevalence to be significantly higher. 5.2.2 Affected communities 'hard to reach' The study states that several respondents considered that communities typically affected by forced marriage (mainly South Asian Muslim communities) were harder to reach as they were 'inward-looking, heavily reliant on community-based services operating through local mosques, and generally mistrusting of statutory agencies such as Social Care and the Police' (p37). Factors which contributed to the communities being 'hard to reach' included: - Services being unable to work with children without the full consent of their parents. - Difficulties in getting young women to attend appointments in the office. - Difficulties gaining access to the family home. 5.2.3 Forced marriage detection not a priority The study reports that several respondents said that forced marriage was not a priority in the context of high numbers of domestic violence referrals and generally 'stretched' (p38) children's services. 5.2.4 Forced marriage as a politically and culturally sensitive issue The study suggests that fears of being 'culturally insensitive', by aligning forced marriage with particular communities, were greater amongst statutory than voluntary services.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		5.2.5 Lack of professional understanding of forced marriage Respondents to the study identified a need for better understanding of forced marriage, particularly how it differs from the cultural practice of arranged marriage, and how to recognise and response to forced marriage. Respondents also thought professionals needed a greater awareness of excessive parental control and the links with forced marriage, such as when 'parents start to see forced marriage as an antidote to their child's perceived bad behaviour (e.g. drug-taking)' (p38). 5.2.6 Language barriers and lack of access to interpretation Lack of access to community interpreters was identified by a range of agencies. 5.2.7 Lack of reporting sites and lack of local 24-hour contact points Respondents thought that there were few local agencies where young people, particularly those with 'limited freedom' (p39) could seek advice, support and protection. The lack of out-of-hours services was also a concern, given that respondents thought that young people were most likely to be flown out of the country for a forced marriage at evenings or weekends. 5.3 Factors facilitating detection 5.3.1 Perception of forced marriage as a clear abuse of young people's right to choose who they marry Professionals from the voluntary sector tended to be clearer about when cases should be categorised as forced marriage than those from the statutory sector.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		5.3.2 Empowerment of young people through information about their rights Respondents thought that awareness work in schools and colleges was vitally important to increase detection and reporting of forced marriage. 5.3.3 Raising awareness of forced marriage among teachers, learning mentors and personal advisors In addition to raising awareness of young people, respondents thought it was important to raise awareness amongst teachers and other education professionals. 5.3.4 Multi-agency forced marriage training Several respondents had found multi-agency training on forced marriage helpful, for example sessions by the Forced Marriage Unit. 5.3.5 A focus on listening, signposting and protection services Respondents noted that many young people did not want formal statutory interventions, and that advice and guidance for the young person were often sufficient. 5.3.6 Information-sharing protocols between agencies Respondents stated that detection of FM was supported by good information-sharing between agencies, for example between police and domestic violence teams. 5.3.7 Using direct methods of communication young people	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Due to the risks for some young people of meeting a professional face to face, especially in their home, respondents talked about usefulness of communicating via mobile phone or text, sometimes using these to set up face to face communication in schools and colleges.	

11. Liao LM, Elliott C, Ahmed F et al. (2013) Adult recall of childhood female genital cutting and perceptions of its effects: A pilot study for service improvement and research feasibility. Journal of Obstetrics and Gynaecology 33: 292–5

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim Aim was to: 1) Explore women's recall of FGM and perceptions of long term consequences; 2) Explore feasibility of future research. **Methodology** • Qualitative study. Interview study. **Country** • UK. Although participants were Somali speaking women, assume that some of their experiences had occurred outside the UK, although this is not explicitly reported in the paper.	**Participants** • Adult survivors of child abuse - 17 Somali speaking women. **Sample characteristics** • Age - Not reported. • Sex - All female. • Ethnicity - 'Somali speaking' - ethnicity not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - 7/17 (41%) reported mental health problems due to FGM. 7/17 (41%) attributed some current physical health problems to FGM. • Sexual orientation - Not reported. • Socioeconomic position -	**Other/general** **• Narrative findings** Recall of FGM. The majority of women had undergone FGM in early childhood: 3–5 years 7/17, 6–8 years 7/17, 9–12 years 1/17, 13–16 years 2/17. The study reports that none of the women consented to FGM. In terms of who had wanted them to have the procedure women reported this as: their mother 8/17, their parents 5/17, a female relative 3/17. The study reports that some participants were aware of individuals who did not want them to have the procedure: their mother 2/17, their father 6/17, siblings 2/17, their doctor 1/17. Participants reported that the procedure was carried out by: a doctor 6/17, a friend 4/17, a traditional woman 4/17, a pharmacist 1/17. Eight of the 17 women reported memories of pain in relation to the procedure, 2/17 reported having received an offering. Perceived long term effects of FGM. Seven out of 17 participants said they had current physical health problems as a result of the FGM, 4/17 reported sexual difficulties, 2/17 reported fear of men,	**Overall assessment of external validity** - **Overall assessment of credibility (internal validity)** + **Overall score** - Lack of information regarding where FGM was conducted is a significant omission in terms of us being able to draw conclusions from this study relevant to our review.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Source of funding • Other - Appears to have been funded by UCL Hospitals NHS Foundation Trust, with additional financial assistance from Ampelos Trust.	2/17 (12%) women reported university education. 1/16 (6%) had elementary education. Other participants 'somewhere in between'. • Type of abuse - 7/17 (41%) had Type III FGM; 3/17 (18%) Type II; 1/17 (6%) Type I; 3/17 (18%) another type of FGM; 3/17 (18%) not examined. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** n=17.	4/17 reported emotional or family problems. One woman said her life had been 'completely blighted' by FGM. Two out of 17 said that overall FGM had affected their lives negatively. Attitudes towards FGM. Fourteen of the 17 women agreed that FGM should be eradicated. All women disagreed that adults should have a right to expect their children to undergo FGM. Eleven out of 17 disagreed that doctors should re-infibulate women.	

12. McElvaney R, Greene S, Hogan D (2014) To tell or not to tell? Factors influencing young people's informal disclosures of child sexual abuse. Journal of Interpersonal Violence 29(5): 928–47

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The aim of the study is to focus on 'to better understand the factors influencing informal disclosure of child sexual abuse experiences' (p928). Informal disclosure is defined as disclosing to a family member or friend.	**Participants** Children and young people - 22 young people who had experienced child sexual abuse Caregivers and families - 14 parents of these interviewed young people who had experienced child sexual abuse. **Sample characteristics** • Age - The majority of the young people who were interviewed for	**Narrative findings –** The findings are explored in five key domains: 1. 'Being believed' – The most shared theme was fear of not being believed (n=14). Participants commented on reasons behind this, for example one 16 year old girl expressed self-doubt: 'I had to deal with it for a long long time and I dunno … the more you leave it unsaid the more unbelievable it becomes'. Findings suggest that the fear of not being believed was unfounded because when most of the young people did	**Overall assessment of external validity** ++ Overall, study meets most of the quality criteria, however caution to generalise the UK as the study is based in a child sexual abuse assessment and therapy centre in a hospital in Ireland.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Methodology: Qualitative study. The use of in-depth semi-structured qualitative interviews were conducted with 22 young people who had experienced child sexual abuse and 14 parents of these young people. Methodology was informed by grounded theory and the authors provide a summary of the data collection and analysis which was informed by methods developed by Hill et al. (1997). **Country:** Not UK. The study was carried out in Dublin, Ireland. **Source of funding:** Government - The author disclosed receipt of the financial support from the Health Research Board, Ireland.	the study were between 13 and 18 years (n=20). The other 2 young people were between 7–12 years old. It is not stated how old the parents of these children were. • Sex - Of the 22 young people interviewed, 16 were female and 6 male. Of the 14 parents interviewed, 12 were female and 2 were male. The majority of parents interviewed were mothers (n=11), and 1 was a father only. A set of parents was all interviewed together. • Ethnicity – Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - The type of abuse explored in case study relates to sexual abuse. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - N/A. **Sample size** The study explores 22 young people who had experienced child sexual	disclose, they were believed. Parents who were interviewed describe how they were sceptical of their child's disclosure remarking that they perceived 'her daughter had misinterpreted an inappropriate touch from her partner' (p934). After initial reactions of disbelief, parents then comprehended and believed the disclosure. 2. 'Being asked' – The study found that young people were asked in different ways if they were being abused and in some instances, there were behavioural changes in the abused young person. Eleven young people described being asked 'explicitly if they had been abused' (p935), where other young people recalled being asked what was wrong, some by their family, friends or partner just knowing and the disclosure being made after probing. Young people stated that their behaviour led to disclosure for example, through self-harming or not eating which then drew them to the attention of professionals, i.e. counsellor or youth leader. Parents described seeing sexualised behaviour between 2 young people which then led to disclosure however conversely, 1 parent commented that they had no suspicion that their child was being abused. 3. 'Shame/Self-blame' – A common theme expressed by over half of young people (n=16) was feeling ashamed and guilty of the abuse. Various reasons were voiced by the young people for why they felt ashamed: where they felt too embarrassed to discuss the abuse with their parents; not fighting back against the abuse; and the abuser told them they were to blame. The study highlighted self-blame as a subsequence issue as years progressed, 1 young person	**6. Overall assessment of credibility (internal validity)** + The study explores the justification of adopting the methodological technique, however there is little consideration for the limitations and generalisability of the study findings. **Overall score** ++ A good, thorough empirical study which meets its research aim.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	abuse and 14 of these young peoples' parents' perspective on disclosing.	expressed that 'I must be a certain type of person' (p936). In one instance, a young person retracted their statement because they felt guilty about their siblings missing the father (perpetrator) and couldn't understand why their father did not live with them any more. On the contrary, one 13 year old girl understood that she was not to blame or at fault. 4. 'Fears and concerns for self and others' – Fear ranged from feeling afraid during the experience of the abuse; feeling afraid of telling; and being scared of the consequences. Consequences that the young people described were to do with fear that they would break their family up, get their abuser into trouble that they themselves might would get into trouble, and what people would think of them. Interestingly, 1 young person remembered a fear of the legal proceedings, 'I had an awful fear about standing up in court in front of him' (p937). The study found that some of the young people's fears did not materialise, whereas others were not unfounded, in that they did upset their family, i.e. taking the dad away from their siblings. Fear of abuse against sibling or other children was also raised as a concern. Parents that were interviewed described their reactions upon hearing that their child was abused of evoking hysteria and feeling that 'it was the end of the world'. 5. 'Peer influence' – 15 young people discussed that initially they disclosed the abuse to a peer, be it friend, boyfriend, or cousin before telling an adult about the abuse. Young people recalled that this led to them being encouraged to tell an adult and highlight that it was 'a very bad situation' (p939).	

13. McNaughton Nicholls C, Harvey S, Paskell C (2014) Gendered perceptions: what professionals say about the sexual exploitation of boys and young men in the UK. London: Barnardo's

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The study has 4 research questions, 1 of which matches our review question which is: - to 'suggest ways in which policy and practice may be able to identify and appropriately respond to male victims of CSE, as well as those at risk' (p13). The other 3 questions are less relevant to this review question which are: - identify perpetration and victimisation processes apparent in male-victim CSE cases known to professionals - explore existing service provision for boys and young men at risk of or experiencing CSE - identify future research priorities (p13). **Methodology:** Qualitative study. This paper reports a qualitative study. It appears that this was undertaken as part of a wider study (summary	**Participants** Professionals/practitioners -Professionals with experience of working with boys and young men experiencing, or at risk of, sexual exploitation. **Sample characteristics** • Age - Not reported. • Sex - Female: n=29 Male: n=21 • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** n=50, comprising 41 qualitative interviews and 9 'online responses' - unclear what the online responses involved.	**Narrative findings** Findings, reported from Section 3, will explore professional practice in relation to boys and young men at risk of sexual exploitation. 3.1 Identification 3.1.1 General barriers to disclosure Professionals interviewed reported a number of barriers to young people disclosing sexual exploitation which applied to all young people of any gender. These were: - Fear of professionals' responses: e.g., that they would not be believed, that people would think they had consented, or that their information would not be kept confidential - Fear of perpetrators' responses: e.g., fear of retaliation by perpetrators, or that they would encourage others not to believe the young person - Grooming processes: e.g., young people may not perceive themselves to be victims, may not want to lose the 'benefits' of the exploitative relationship, or may fear revealing further sexual abuse within their family. Professionals also identified the following groups as being at particular risk: - Children with learning disabilities who may lack capacity to understand they are being exploited, or not trust their own recollections - Children from BME backgrounds, who may have additional fears of bringing dishonour/shame to their families (although noting that other research has shown that BME young people accessing CSE support services is roughly proportionate) (Cockbain et al. 2014). 3.1.2 Barriers specific to gender	**Overall assessment of external validity** + UK study but only part of overall research aim was relevant to our review question. **6. Overall assessment of credibility (internal validity)** + **Overall score** + Only part of overall research aim was relevant to our review question. Study is of reasonable quality, although limited exploration of divergent perspectives across different types of interviewees.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
reported in McNaughton Nicholls et al. 2014 'Research on the sexual exploitation of boys and young men'). **Country:** UK, England. **Source of funding:** Voluntary/charity - The Nuffield Foundation.		The study reports that professionals thought that boys and young men were less likely to disclose sexual exploitation than young women. This meant they relied more on information from other professionals or the young person's friends. One professional said: 'The boys that we're working with don't disclose. So ... we're getting an idea from the various professionals [about whether a boy is at risk of exploitation] ... but unfortunately we are filling the blanks with assumptions' (Service manager, CSE service) (p30). Professionals also reported that the friends of girls who are being exploited are more likely to report it than the friends of boys. The study reports the following gender-specific barriers to disclosure for boys and young men (taken from Table 3.1, p31): 1. Discrimination, social attitudes and stereotypes - homo/bi-phobia and trans-phobia; stereotypes of masculinity; stigmatisation of boys and young men as offenders. The study reports that professionals said that young men who were being sexually exploited by men were afraid of experiencing homophobia from professionals, as well as their friends and the wider community. Professionals acknowledged that young women might also fear homophobia, but thought the fear was greater amongst young men, as it was linked to 'perceptions of socially acceptable masculinity' (p32). Professionals thought that young men found it hard to identify themselves as 'victims' because this did not fit with the stereotype that men should be able to look after themselves. Professionals also reported a lack of trust in statutory services in cases where young people had had criminal involvement in the past, and were worried that they would be perceived	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		as becoming abusers themselves. One professional said: 'The other thing is, having worked with a lot of male survivors, they have said to me that they have never wanted to disclose to any statutory agency, because they're frightened they would be viewed as an abuser, so they won't tell their GP, they won't tell anybody, because they're terrified, and that's a very difference experience for females as well' (CSE policy specialist) (p32). Professionals also thought that trans* young people would particularly fear disclosing abuse, because their gender identity might be seen as a reason for the abuse. 2. Gender differences in CSE education - Lack of male-victim-focused CSE education; lack of knowledge of support services. Professionals also talked about CSE education, commenting that this is predominantly targeted at young women. They thought this could also have an impact on young men's willingness to disclose to professionals. They thought the lack of CSE education targeted at boys meant they may not have the "language" to recognise or talk about themselves as victims' (p32), and were also less likely to be aware of support services. 3. Gender differences in emotional responses - Emotional isolation; weaker communication skills than girls and young women; desire to move on. The study reports that professionals described gender differences in emotional responses to exploitation between young men and young women. Professionals reported that young men were often more 'emotionally isolated' (p33), and less likely to talk about their feelings. Professionals also thought that young men were more	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		likely to wish to 'move on' (p32) rather than seek support. 3.2 Professional practice and the identification of male victimisation Professionals noted the following professional barriers to identification of exploited boys and young men (taken from Table 3.2, p34). 1. Discriminatory social attitudes and stereotypes - Poor understanding of sexual identities; belief that young men do not need protecting; boys and young men viewed as offenders. 2. Gendered implementation of identification practice - Boys scoring lower on risk assessments than girls; gendered interpretations of indicators. Professionals noted that there may be assumptions that a young man is gay because he is being sexually exploited by a male perpetrator. Professionals also reported that, in some cases, boys and young men who identified as gay or were questioning their sexuality were automatically assumed to be at risk of CSE. Professionals also noted their own stereotypes of masculinity may play a role. One said: 'The instinct to protect boys is not there because they should protect themselves. It's almost that that we're fighting against' (Service manager, CSE service) (p34). Furthermore, professionals thought that, whereas girls were thought of as victims, boys were often considered to be offenders, with service responses focusing on criminal behaviour. They thought that boys were more likely to be identified only after exploitation had occurred, through receiving services as a result of offending behaviour.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		3.2.1 Gender stereotypes and professional practice Professionals thought that indicators of CSE were more likely to be perceived as applying to young women. For example, professionals may be less likely to think of young men as being vulnerable when intoxicated. Professionals also reported that young men were less likely to receive a CSE assessment where indicators were observed, and if they did receive one, were likely to be assessed as at lower risk. One professional said: 'And it's quite interesting because there was a case study given to us of a 19-year-old taking pictures of a 15-year-old and then, if that was a 19-year-old man taking indecent images of a 15-year-old girl we'd all be concerned and it would be a child protection issue. But because it was a 19-year-old girl taking pictures of a 15-year-old boy, everybody's very blasé about it. And you know, lucky him and he's having the time of his life, sowing his wild oats, getting his experience. A whole host of things. The poor boy's in exactly the same position as the young woman, and we don't see that' (CSE policy specialist) (p35). 3.3 Promising practices in identification Professionals reported three forms of practice which they thought were effective in improving identification of boys and young men at risk of CSE: - gender-neutral materials - providing training for professionals on male victims - co-location of CSE specialist practitioners with statutory agencies.	

14. NSPCC (2013) Would they actually have believed me? A focus group exploration of the underreporting of crimes by Jimmy Savile. London: NSPCC

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To seek information from victims of Jimmy Savile about what had prevented them from reporting to the police at the time of the abuse, and 'to explore how police can improve their management of the reporting process and subsequent interviews and contacts' (p4). **Methodology:** Qualitative study. The research consisted of five focus groups attended by a total of 26 people who had been abused by Jimmy Savile. **Country:** UK. Focus groups held in London (2 groups), Leicester, Liverpool and Leeds. **Source of funding:** Government - Her Majesty's Inspectorate of Constabulary.	**Participants** Adult survivors of child abuse. We have assumed that the participants were adults at the time of the study, although this is not made clear in the methodology. The study notes that 4 participants had been adults when they were abused. We have decided to include the study nonetheless as this implies that the majority of participants were children when they were abused. **Sample characteristics** • Age – Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** n=26.	**Narrative findings** The findings of the focus groups are reported in response to each of the questions posed. Q1. At the time of the abuse, how aware were you that what had happened to you was abusive? The study reports that some participants were aware that they were being abused, whereas others reported that although they felt 'uncomfortable' and 'frightened' they weren't fully aware of what was happening to them. Some participants reported that they had felt at fault, or wondered if they should feel grateful that a celebrity had 'chosen them'. Q2. Did you tell anyone about the abuse at the time? The study reports that a minority of participants told someone about the abuse at the time. People who were told included extended family or friends or hospital staff. One person had reported it to the police, but it was unclear what action resulted. The reason given by most people for not telling was that they thought they would not be believed, particularly given that Jimmy Savile was a 'powerful and influential' adult. Participants reported feelings of guilt, shock, embarrassment and shame. For participants who had been living in residential care at the time of the incident, they reported that they had been concerned about how a disclosure might affect decisions about their care. Q3. In recent years, prior to the media coverage, have you spoken to anyone about the abuse? The	**Overall assessment of external validity** - Lack of consideration of ethical issues, and of transferability of findings given the very particular circumstances involving a high profile celebrity and subsequent documentary film raising awareness. **Overall assessment of credibility (internal validity)** - **Overall score** - No consideration of ethical issues reported. Little consideration of transferability of the findings to other cases of abuse, given the particular circumstances (i.e. high profile celebrity case). Little detail given regarding participants, or methods of analysis.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		study reports that some participants said they had discussed the abuse with, for example, a partner, extended family or close friends. None of the participants had gone to the police or other authorities. Q4. Although you did not contact the police at the time of the abuse, did it occur to you that you could report to them directly? The study found that the people who were adults at the time of the abuse knew that they could report to the police, but decided not to. The participants who had been children at the time of the abuse did not know that they could go to the police. One participant said: 'I never thought I could go to the police on my own ... children's minds work completely differently, don't they?' Q5. What enabled or encouraged you to report abuse at this time? Participants reported that the media coverage of allegations against Jimmy Savile, particularly the stories of other victims, had encouraged them to report abuse. Participants also noted the importance of support from family and friends. Reasons given for reporting the abuse following the media coverage included: - To acknowledge the impact the abuse had had on their lives, and to get a sense of 'closure' (p10) - To 'support and corroborate the experiences of those who had spoken out in the documentary and were not believed. Participants thought they would have reported the abuse sooner if there had been earlier media attention, or if they had known that there were other victims. Q6. How did you report your concerns? Participants had reported their concerns to the police, the makers	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		of the documentary about Jimmy Savile, ITV, the NSPCC Helpline or a solicitor. Q7. How affected were you by the media coverage? The study reports that participants reported a number of reactions, including feelings of anger, flashbacks or feeling physically sick. However, the negative impact of the coverage was thought to be balanced out by encouraging people to disclose the abuse. Q8. What was your experience of reporting the abuse now? What support were you offered afterwards? The majority of participants had had positive experiences of liaising with the police. People's perceptions of the police appeared to depend on the extent to which the police were formally responding to their disclosure. Most participants reported having to repeat their disclosure a number of times. Q9. Do you think there are any barriers to reporting abuse now? Participants thought that being a 'lone voice' (p15) might still impede people from coming forward to disclose abuse. Participants thought there should be a specialised channel for people disclosing sexual abuse. Q10. What changes could be made by the police to encourage people to report abuse at the time it occurs? Participants identified the following barriers to disclosure including: - Police not so much part of the community - Limited access to local police stations - inaccessible locations and limited opening times. Participants thought the police should do more in schools to inform children and young people about the role of	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		the police. They also thought that officers investigating claims of sexual abuse should have specialist training. Q11. Mandatory reporting - If other professionals were given a legal obligation to report all allegations of abuse that they became aware of, would this improve the protection of children? The focus groups explored the idea of mandatory reporting. All participants thought that it would help children if professionals had to report concerns or allegations of abuse. However, the study reports that 'concern was expressed about the idea of criminalising professionals who did not report allegations, and also that this may lead to professionals becoming very anxious and almost reluctant to engage with young people in the same way' (p17). Q12. Advisory capacity of police - If victims were able to contact police officers to discuss their abuse and seek advice, without the fear that this would automatically be reported as a crime and acted upon, would this encourage victims to interact with the police? Most participants thought this would be a good development, and that this should be provided by a single nationally recognised service. Wider context - impact of abuse. Most participants thought the abuse they experienced had negatively impacted on them. For example, participants referred to drug and alcohol misuse, risk-taking behaviours, running away, anti-social behaviour and poor relationships with parents and carers. One participant said: 'I became very withdrawn ...They took me to the doctor and put me on anti-depressants and I've been on and off them ever since' (p19). Many of these problems continued in to	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		adulthood, particularly drug and alcohol dependency, depression, suicidal thoughts and relationship problems. None of the participants remembered anyone asking them about the changes in their behaviour when they were a child. They thought that if someone had spent time and listened, it may have encouraged them to talk about the abuse. Some participants said their experience had left them with a sense of 'blurred boundaries', meaning that they were vulnerable to abuse by other adults.	

15. Pearce J, Hynes P, Bovarnick S (2009) Breaking the wall of silence: practitioners' responses to trafficked children and young people. London: NSPCC

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The aims of the research are as follows: 1. Explore in depth the different ways in which trafficking is understood by a range of practitioners from different service agencies and provide evidenced recommendations for practice in their area. 2. Explore the obstacles that might emerge to identifying the numbers of young people trafficked in the three areas.	**Participants** Children and young people. Reviewed case files of 37 trafficked children and young people. Professionals/practitioners - 65 practitioners involved via focus group, a subset of these were also interviewed (numbers not given). An additional 7 practitioners were interviewed only. **Sample characteristics** • Age - Practitioners in focus groups/interviews: Age not reported. Young people's case files: Age 3 and under n=6, age 4–8 n=1, age 9–12 n=1, age 13–15 n=15, age 16–17 n=14.	**Narrative findings** 1. Identification of trafficking – 'trafficking is a process, not an event' The study reports that practitioners noted that identification and disclosure trafficking are rarely one-off events, but an ongoing process which requires the building of trusted relationships. The study notes the following facilitators to identification and helping: - Not allowing age or immigration status concerns to override child protection concerns, which should be paramount. - Not assuming that an interpreter from the same community is the best choice, when in fact they may represent to the child the community which has exploited them.	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** ++ **Overall validity rating:** ++ Thorough data collection, analysis and reporting.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
3. Identify the numbers of children and young people trafficked into each of the three areas. 4. Chart the process through which a child or young person first gained access to a support agency, including how they first contacted an agency and for what reason. 5. Where possible, provide a profile on each of the children and young people identified including: age; nationality; country of origin; the reason they were trafficked into the country; and a summary of their current circumstances. 6. Identify how the practitioner understood the immediate and longer-term needs of the children and young people concerned. 7. Identify how the professionals feel these needs are best met. 8. Where possible, identify perceptions of	• Sex - Practitioners in focus groups/interviews: Gender not reported. Young people's case files: Girls n=30, Boys n=4, Gender not known n=3. • Ethnicity - Practitioners in focus groups/interviews: Ethnicity not reported. Young people's case files: Ethnicity not reported, but information on nationality provided. Country of origin: UK n=1-, China n=8, Nigeria n=8, Somalia n=1, Pakistan n=1, Cameroon n=1, Ghana n=1, Congo n=1, Sierra Leone n=1, Zimbabwe n=1, Uganda n=1, Eastern European Country n=1, Unknown n=1. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Of the case files examined, 10 related to cases of trafficking of UK citizens and 27 to cases of trafficking in to the UK from abroad. Reasons for trafficking were as follows: Sexual exploitation n=19 (this included 9 of the trafficked UK citizen), benefit fraud/illegal adoption n=7, domestic servitude n=5, forced marriage n=2, restaurant work n=2, drug trafficking n=1, not known n=1.	- Having continuity with the same interpreter, keyworker or legal guardian. - Use of an independent guardian. - Recognising the trafficked children may not recognise themselves as having been abused, or having been deprived of a 'childhood' as their understandings of childhood and home may differ from those in the UK. One practitioner said: 'Some of the Chinese boys who are over 15 would think that why shouldn't I be working? I have come here to get a better life and, yes, money has exchanged hands for me but I know it's illegal but what is the problem?' (Interview 10, p67). - Not allowing the image of trafficking for sexual exploitation to 'overshadow' (p6 executive summary) awareness of the other forms of exploitation, including benefit fraud, forced marriage, domestic servitude or work in cannabis factories or nail parlours. Also, remembering that children who are originally from the UK can be trafficked, and that both girls and boys can be trafficked. - Recognising the children trafficked from outside the UK face specific problems, which can include language barriers; experiences of war, famine or poverty; insecure legal status and unfamiliarity with UK cultures and systems. - Actively working to keep track of, and support, young people who go missing. The study also notes that: - Practitioners may be unsure how to apply a definition of trafficking in their practice, and also whether a young person had been trafficked or smuggled and 'whether the distinction mattered in practice' (p60).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
how the children/young people feel these needs are best met. 9. Make recommendations about how agencies or individuals can best support the children/young people concerned. We considered question 2 to be relevant to our review question on Recognition, and 7 and 8 to be relevant to our review question on Response. **Methodology:** Qualitative study. - Focus groups with 65 practitioners. - Interviews with a selection of focus group practitioners (number not specified) and with an additional 7 practitioners - Case file analysis of cases of 37 trafficked children and young people. **Country:** UK, England. **Source of funding:** Voluntary/charity -	• Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - All young people were trafficked. **Sample size** Practitioners: n=72 Children's case files: n=37.	- Trafficking can be 'hidden' within private foster care arrangements (p64). 2. Disclosure of trafficking – 'trafficking can be hidden behind a wall of silence' (p7 executive summary) The study reports that: - Both interviews and case file review showed that traffickers manipulate young people into being trafficked. This means that young people are likely to be confused and traumatised, and may 'block out' (p7 executive summary) their experiences as a way to cope. - This means that children will only disclose after they have built a relationship of trust. It also means that disclosures are likely to be elicited rather than accidental or purposeful by the young person. - The study further notes that practitioners thought that a number of practitioners from different agencies to interview children might 'in itself, be abusive' (p90). - Practitioners reported that young people's accounts of their experiences may be hard to understand, or contain discrepancies. This can mean that they are not believed. - The process of disclosure was felt to be different for UK nationals compared to those trafficked from abroad, but still complex and often characterised by threats and pressure from traffickers. The study also notes that disclosure can be more difficult for boys and men. - Practitioners felt that their practice had improved through experience.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study reports that research has been funded by 'The Children's Charity' (p6) (unclear if this refers to NSPCC or another charity).		- Practitioners thought they should be encouraged to understand the full variety and complexity of trafficking, rather than reducing understanding to particular categories or profiles.	

16. Rees G, Gorin S, Jobe A et al. (2010). Safeguarding young people: Responding to young people 11 to 17 who are maltreated. London: The Children's Society

Research aims.	PICO (population, intervention, comparison, outcomes).	Findings.	Overall validity rating.
Study aim: The aim of this study is to explore 'access to, and initial responses of, services for young people with potential maltreatment ... to promote protective responses for this target group' (p7). The section relevant to this review question is entitled young peoples' experiences of seeking help. **Methodology:** Qualitative study. The relevant methodology relating to this scope involved in-depth interviews with 24 young people who had been	**Participants** Children and young people. The 24 young people who were interviewed for our study either had social care intervention from an early age or had first come to the attention of Children's Social Care Services in between the ages of 11 and 18. The study includes 14 boys and 10 girls, and the majority of participants were White British majority (n=18). One young person was British Asian and the study included unaccompanied asylum seeking children who were originally from Afghanistan (n=3) and Eritrea (n=2). **Sample characteristics** • Age - Young people ranged between 11 and 18, with categories determined: 11–14 (n=5); 15–16	**Narrative findings** The findings provide a narrative summary of young people's experience namely: (1) the difficulties with seeking help; (2) seeking help from peers; (3) seeking help from family members; (4) seeking help from professionals. 1. Difficulties with seeking help: Many young people reported that they were concerned about not being believed or not knowing who to tell. There were fears expressed relating to whether they would be put into local authority care so young people were not always direct with telling a professional. One young person discussed colluding with her abusive mother to disguise from the social worker what was happening: 'I used to have bruises, the lot, and we just used to make up stories' (Anna, age 17; p43). Conversely, other young people actively sought placement outside of home. Some young people described being concerned about consequences which echoed previous research as a barrier to disclosure (Baginsky 2001).	**Overall assessment of external validity** ++ Study relates to question of exploring young peoples' views and experiences of recognition. **Overall assessment of credibility (internal validity)** + (When taking in to account additional info from Jobe and Gorin.) The study does not have a rigorous methodology or consideration of limitations. Presentation of infor-

Research aims.	PICO (population, intervention, comparison, outcomes).	Findings.	Overall validity rating.
referred to children's social care aged 11–17. This study is also reported in Jobe and Gorin (2013). This paper reports a briefer version of the study findings, but has more detail on study methods. Where necessary, additional methodological information has been taken from Jobe and Gorin (2013). **Country:** UK. **Source of funding:** Not reported.	(n=13); and 17–18 (n=6). Sex -The study includes 14 males and 10 females. • Ethnicity -Participants were White British majority (n=18). One young person was British Asian and the study included unaccompanied asylum seeking children who were originally from Afghanistan (n=3) and Eritrea (n=2). • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - The maltreatment experience was divided into 2 groups – those who had suffered maltreatment from an early age therefore were receiving social work intervention from an early age but were still receiving support between the population age of 11–17 (n=6). The majority came to the attention of children's social care (n=18). Reason for referral included a range of issues such as 'homelessness, being thrown out of home, mental-health problems, alcohol and drug misuse, behavioural problems, risk-taking	Young people commented that their abusive parent had prevented them from seeking outside help, where in 1 instance a young person was barred from attending school. Findings suggest that it is essential, that young people need to have an established relationship with a professional (a teacher or a youth worker) and require 'confidence and safety' (Emma, age 14; p44). 2. Seeking help from their peers: Data suggests that young people generally sought help from their peers about the abuse they were experiencing rather than a professional or family member. This support was valued highly and the findings suggest targeting information to young people to know what constitutes abuse and where to seek help. The study suggests that young people did not experience difficulty in disclosing abuse to their peers, but in some instances, abuse was used as a way of bullying the young person. 3. Seeking help from their family members: A number of young people disclosed to a parent or family member, however where the parent was the perpetrator, extended family were asked for help. The study found that generally these young people were not believed which in turn discouraged them from seeking help outside of the family. Some young people were referred to children's social care services after disclosing to a family member, and usually, but not always, if the perpetrator was independent of the family. For example, 1 young person said: '[I said to my auntie] that my dad hits me and stuff like that. But she just said that - well, she couldn't do anything because, like,	mation is difficult to ascertain where data is collected making conclusions challenging to draw. In addition, there is discrepancy in young people's age as referred to in text as both: 11–17; and 11–18. **Overall score** + The study is suitable for scope and the findings enrich discussion about barriers to young people disclosing sexual abuse. Drawing on additional information from Jobe and Gorin (2013), where the research design is more informed, the findings are more convincing as data is richer and analysis is clearer.

Research aims.	PICO (population, intervention, comparison, outcomes).	Findings.	Overall validity rating.
	behaviour, violence and conflict with parents' (p39). • Unaccompanied asylum seeking, refugee or trafficked children – The study included unaccompanied asylum seeking children who were originally from Afghanistan (n=3) and Eritrea (n=2). **Sample size** n=24.	family and stuff. So she just told me to stay out of his way' (Fatima, age 15) (p46). 4. Seeking help from professionals: The study found that generally young people sought help from their school teacher, although 1 young person contacted the police and the 5 unaccompanied asylum seeking children interviewed were referred to children's social care through the police. Some young people discussed feeling unsure of who to disclose the abuse to and were confused over the roles of professionals. The study discussed more in-depth young people's experiences of seeking help from both their teacher and the police. When young people did seek help from their teacher, their experience had been positive about the established and valued relationship, and the help they received. However, in some instances, young people expressed that they felt their disclosure had not been progressed or taken seriously which left young people feeling unsupported. One young person described their teacher 'like sometimes like they listen but they're not really listening' (Laura, age 15; p45). When 1 young people discussed their experience with the police, they explained that the experience had been daunting. One young person said of their experience with the police: 'I don't know – I suppose they need to be sort of not as dismissive with young people. I suppose – when we first went into the station I got the feeling we were looked down on … The people at the reception weren't – they were very – I don't know- I suppose I got the feeling because we were young – young, youths and that, that they thought we were in trouble but it wasn't like that and it felt like they dismissed us a little bit and that when we were in	

Research aims.	PICO (population, intervention, comparison, outcomes).	Findings.	Overall validity rating.
		the waiting room, but the woman we spoke to was really nice. I suppose they should be more welcoming and have more people on hand at police stations and things like that specifically for young people cos when I first went and I spoke to someone who I don't think had anything to do with child protection or anything like that. Spoke to someone completely different who then referred it over. So maybe if there was more people, people who were aimed at talking to younger people then people would feel more able to sort of speak out and come forward with things like that' (Lisa, age 15; p48). The 5 unaccompanied asylum seeking children are described to have not directly spoke with the police but interpreters and most described the experience with police positively. However, 1 young person discussed the experience negatively. He said: 'I mean they was not like helpful. Like when we come to this country … I speak several language, I can speak seven, eight languages … and they were really swearing to us, and they were very bad at that time, but they didn't know that I can understand them. But I didn't say nothing to them because I was very scared that time because we don't know what will happen to us' (Khalid, age 18; p50).	

17. Rigby P (2011) Separated and trafficked children: The challenges for child protection professionals. Child Abuse Review 20: 324–40

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To scope the prevalence of child trafficking, profile children and identify factors that facilitate or hinder intervention.	**Participants** Children and young people. Case files of 75 unaccompanied asylum-seeking children. Professionals/practitioners - 16 front-	**Narrative findings** Factors that hinder recognition Cultural barriers	**Overall assessment of external validity** + Assessment is a subset of overall research question.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Methodology: Qualitative study. The use of qualitative interviews were conducted with 16 experienced frontline professionals: 7 individual interviews, and two focus groups. Grounded theory was adopted to facilitate practitioners to explore their experiences of trafficking. Interviews were transcribed and manually coded to sort into themes. For accuracy, participants were able to review their transcript, contribute to the final report and make amendments where appropriate. **Country:** UK, Scotland (Glasgow). **Source of funding:** Not reported.	line professionals working with separated or trafficked for children for >3 years. **Sample characteristics** • Age - Children and young people aged between 12 and 17. Professionals: Not reported. • Sex - Children and young people: 38 females and 37 males Professionals: Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation – Not reported. • Socioeconomic position - Not reported. • Type of abuse - 16 (21%) of the unaccompanied asylum-seeking children in the case file sample were categorised as having been trafficked. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - All children in case file sample (n=75) were unaccompanied asylum-seeking children. 21% were deemed to have been trafficked.	- Working with young people from various countries is a challenge, as one professional recalls, 'How do you work with young people from 23 countries ... people can't tell you' (p329). - Lack of understanding about 'potential victims of trafficking may be ... exacerbated by professionals lack of understanding of what trafficking is and how children's experiences before, during and after movement affect them' (p329). - Young people might not think they are being exploited. Trauma and fear - Professionals comment that most, if not all, children they work with experience 'absolute fear of everything' (p330), therefore some young people were seen not to be forthcoming with their stories. Trafficker-child relationship - Some professionals noted that there were a few examples where young people would maintain the relationship with their trafficker and in some instances, young people felt they were benefiting from the relationship. Child-professional relationship - As noted, 'a prerequisite for effective practice was the development of maintenance of a trusting working relationship' (p332). Some professionals in the study reported struggling to develop a positive relationship because they experienced young people who were scared and confused over their role. Identification and assessment	**6. Overall assessment of credibility (internal validity)** - **Overall score** - Not clear which data were gathered via interview, and which via focus group. Analysis methods unclear. Relatively little reference to, or presentation of, primary data gathered.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Sample size 75 unaccompanied asylum-seeking children. 16 frontline professionals.	- Background and journey information is difficult to corroborate - therefore difficult to identify whether trafficked. - There are limited links with international agencies who could help with corroboration. - Variability in whether children will disclose trafficking 'Some of them will tell you ... they will come in traumatised and upset and tell you quite early on. And others just won't' (p333). - Traumatised young people do not always present a coherent story.	

18. Stanley N, Miller P, Richardson Foster H (2012) Engaging with children's and parents' perspectives on domestic violence. Child and Family Social Work 17: 192–201

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: Study is to understand children (n=19) and parents' (both survivor and perpetrators - n=11, 10 respectively) perspective on experiencing domestic violence. The relevance of the study are feelings explored that are barriers to disclosing and acknowledging domestic violence. **Methodology:** Qualitative study. The study interviews 3 groups of	**Participants** Children and young people - There were 5 focus groups held with 19 young participants aged 10–19. Eight were male and 11 were female. The majority (n=16) categorised themselves as white British, 1 as white/Asian, 1 as white/black Caribbean and 1 as white/black African. Caregivers and families - Survivors - 10 of the 11 survivors were female, age range between 25 and 35. Most of the research participants described themselves as belonging to (BME) groups, 4 identified as white British. Between the survivors, they had 26 children.	**Narrative findings** The study includes a narrative account of 'disclosing and acknowledging domestic violence' that is relevant to the scope: Factors that hinder identification: - Most participants commented on the 'stigma, shame and embarrassment associated with disclosing their experience of DV to family, friends and professionals' (p194). One survivor recalls hiding the DV from her family because she was 'ashamed'. Additionally, one young person details stigma experienced by being 'put down at school' (p194). - There was variance between survivors and perpetrators perception in acknowledging the DV impacting their children. The variation occurred because some felt their children were shielded from the abuse but acknowledged that their children might have seen	**Overall assessment of external validity** - The study has relevance in part to exploring young people's perspectives on domestic violence, yet as study has adult perspective too, the voice of the child is represented generically. There is no information about obtaining ethical approval from local authorities. The authors are explicit that the research participants are

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
participants who have experienced DV: young people; survivors; and perpetrators. With young people (n=19), the research team hold 5 focus groups recruited through voluntary and statutory services supporting children affected by DV. Similarly, survivors (n=11) were accessed through a support service but were interviewed separately using semi-structured interview schedule. Perpetrators were also interviewed individually (n=10). All were recorded and transcribed, then were analysed and themed under the interview schedule, however grounded theory principles were adopted to inform the study. **Country:** UK. Research was carried out in two local authorities	Perpetrators - All were male and aged 35–45 (n=10). Six identified as white British, with the other 4 representing BME groups. The perpetrators were selected because the majority had children generally still in contact with their children.	tragic occurrences within the family home, i.e. removal of perpetrator. Whereas other survivors and perpetrators recognised that their children had witnessed DV, which supports a number of young people who admitted to seeing incidents of DV directly. - Some survivors reported experiencing disbelief from professionals which contributed to their helpfulness feeling and unwillingness to seek help in the future. Factors that support identification: -Young people in the focus groups recognised that being listened to, validated accounts and provided information was important. One young person describes a professional who supported her, 'she was really helpful, she spoke to me rather than just my mum, she was the one that gave us the number for the NSPCC' (p196). - Perpetrators, in particular, expressed the need for non-judgmental attitudes from their practitioners.	anonymised and consent was sought. **Overall assessment of credibility (internal validity)** + **Overall score** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
in the North and South of England. **Source of funding:** Voluntary/charity - This study was funded by the NSPCC.			

19. Tucker S (2011) Listening and believing: an examination of young people's perceptions of why they are not believed by professionals when they report abuse and neglect. Children and Society 25: 458–69

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim Aim is to 'deepen debate by offering a detailed and substantial analysis of young people's perceptions and reactions to their treatment at the time of initial disclosure to practitioners' (p458). **Methodology** Telephone interview (n=58); individual interview (n=33); and group interview (n=17). **Country** UK, England. **Source of funding** Not reported.	**Participants** Adult survivors of child abuse - n=102. **Sample characteristics** Age - 18 years old: n=10 19 years old: n=12 20 years old: n=20 21 years old: n=32 22 years old: n=13 23 years old: n=21 Sex - 66=male 42=female. Ethnicity - Not reported. Religion/belief - Not reported. Disability - Not reported. Long term health condition - Not reported. Sexual orientation - Not reported.	**Recognition** **Narrative findings** Factors that hinder recognition are defined into four typologies that occurred in interviews with young people with a percentage ascribed to the amount each factor occurred: Background and baggage category (64%). The individuals background and history i.e. 'baggage', can impact on the capacity for a disclosure. One young person reports: 'Just seeing you as someone needing help it didn't work like that. F***ed up by case notes and meetings and your records from other schools. Like having a criminal record and it felt like that and only 12. That's why I wasn't taken seriously' (p463). - The youth advisory group defined the category to be broadly as: 'Being preoccupied by names, tags, "previous convictions", misdemeanours, things written about you. Saying then that's what you are and that's all you are and that's all you can be. And then not actually believing what you say because of all this' (p463).	**Overall assessment of external validity** ++ Study relates to question of exploring young peoples' views and experiences of disclosure. Sound ethical consideration and the study is co-produced by an advisory group. **Overall assessment of credibility** ++ Good qualitative study with large sample (n=102) and through methodological approach. **Overall score**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Socioeconomic position - Not reported. Type of abuse - Not reported. Looked after or adopted status - Not reported. Unaccompanied asylum seeking, refugee or trafficked children - Not reported.	Family matters category (55%). Defined by the youth advisory group - 'tell your story and have that story believed despite who you are, or where you come from, or what family you belong to' (p464). Reluctance and refusal category (45%). As defined, this category determines the issues professionals experience at the point of the young person disclosing, and the impact this can have on making judgements and assessing risk. Young people talked about 'being able to see the cogs going round' as the practitioner weighed the evidence and appeared to come to a preliminary, or in some cases a more long-lasting, judgement about an allegation (p465). A young person described his experience of disclosing to a professional: 'You could see what was going on the whole time it was like he was looking into you not just looking at you. I did statistics for my degree and the best way I can describe it was he was weighing up the probability, the likelihood. What if I act and it's not true? It might be safer to do nothing if you're not sure. Who's going to believe a kid over an adult? All of that stuff and he actually said as much but used different kinds of words to push it back on me' (p465). Conversely, young people conceded that they understood the difficulties associated with judgement and risk but wanted to reinforce the courage for disclosure to professionals, and that they felt rejected if they were not believed. One young woman discussed 'you brought it on yourself syndrome', which was the response she received due to behaviours such as: being 'tarty', 'sexting it up with the lads', 'acting in your face' and wearing 'revealing stuff' (p466). One young woman e.g. said: 'Boys thought they were easy on with me;	++ Good empirical qualitative study with large sample (n=102) and through methodological approach. The research is relevant to inform young peoples' perception of disclosure.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		anything goes. This boy really, really hurt me physically and I tried to report him … But I got all this "are you sure?", "this is a serious allegation", "I've told you about hanging about lads" stuff and "go away and think about it". I felt like they didn't know what to do for the best … did what was the best for the school and not for me?' (p466). The issue of feeling referred on by a practitioner. One interviewee commented on the response of a practitioner, i.e. 'It's best to contact someone else that deals with this kind of thing' (p466). The youth advisory group defined this category as: 'It includes practitioners not taking risks if those risks are seen to be too great. Sometimes it can include making a judgment based on only superficially listening to what's being said. Some self-protection, some shifting the blame back to the young person, some assumption and false thinking about who a person really is' (p466). Personal relationships (28%). Some young people described abusers exploiting relationships with professionals as a way of masking their intention towards a young person. As defined by the youth advisory group: 'Making clever and deliberate contacts and attachments for your own ends. Getting alongside workers and convincing them you are OK. Using this as a front and a way of making it really hard to make any kind of accusations' (p467).	

Assessment

Review question 7 – What tools support effective assessment of risk and need in relation to child abuse and neglect?

Review question 7 – Critical appraisal tables

1. **Baumann DJ, Law JR, Sheets J et al. (2005). Evaluating the effectiveness of actuarial risk assessment models. Children and Youth Services Review 27(5): 465–90**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: A series of 3 field studies to examine issues surrounding the scientific integrity and practical utility of actuarial models of risk assessment in child welfare. Only data relating to studies 1 and 2 are reported here, as these looked at predictive validity, cross-validated using an objective measure. The third study did not look at examine predictive validity, and so was excluded. Study 1: RCT examining whether use of, or exposure to, an actuarial risk model led to superior judgements relating to substantiation of allegations compared to individuals not using or exposed to the model. Study 2: RCT examining whether use of, or exposure to, an actuarial risk model led to superior judgements relating to re-investigation compared to individuals not using or exposed to the model. **Description of theoretical approach?** No.	**Was the exposure to the intervention and comparison as intended?** Not reported. **Was contamination acceptably low?** Yes. Units, rather than caseworkers, were randomly assigned to conditions to avoid contamination of the treatment from caseworkers in the same units sharing information. **Did either group receive additional interventions or have services provided in a different manner?** Not reported. **Were outcomes relevant?** Yes. Predictive validity measured using substantiated maltreatment (study 1) and repeat investigations (study 2). **Were outcome measures reliable?** Yes.	**Does the study's research question match the review question?** Yes. Evaluating the effectiveness of actuarial risk assessment models. **Has the study dealt appropriately with any ethical concerns?** No. Not reported. **Were service users involved in the study?** No. Case workers as participants. **Is there a clear focus on the guideline topic?** Yes. Testing the accuracy of actuarial risk assessment models in child welfare. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Professionals working with children at risk of, or experiencing, maltreatment.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity rating:** - Reporting of analysis extremely unclear, including whether reported correlations are within-groups correlations between prediction and outcomes, or biserial correlations showing between-groups differences in the accuracy of prediction. Unclear what data points were used for analysis (cases, workers or units) and reported values of n for each condition do not appear to correspond to any of these.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
How was selection bias minimised? Randomised. Study 1: RCT - randomised by unit rather than individual Study 2: RCT - randomised by unit rather than individual. **Was the allocation method followed?** Not reported. **Is blinding an issue in this study?** Blinding not possible. Blinding not possible as allocation to conditions involved use of a different tool. **Did participants reflect target group?** Yes. All participants were caseworkers working with children and families where there were concerns about maltreatment. **Were all participants accounted for at study conclusion?** Not reported.	**Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Not reported. **Was follow-up time meaningful?** Not reported. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Not reported. **Was intention to treat (ITT) analysis conducted?** Not reported. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. **Were the estimates of effect size given or calculable?** No. Pearson correlations reported for some comparisons, but most data reported graphically only. Unclear whether correlations represent within- or between-groups effects.	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Assessment tools. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **Does the study have a UK perspective?** No – USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Were the analytical methods appropriate? Partly. Very unclear description of analytical methods: - It is unclear whether reported correlation data are correlations between prediction and outcome, or biserial correlations comparing predictive validity of different tools. - Unclear what data points were used for analysis (cases, workers or units) and reported values of n for the different groups do not appear to correspond to any of these. **Was the precision of intervention effects given or calculable? Were they meaningful?** Partly. Pearson correlations and p values reported for some comparisons, but most data given graphically only. **Do conclusions match findings?** Partly. It is difficult to judge the extent to which conclusions match findings, as there is very sparse reporting of statistical data, and it is unclear what is being compared within the analysis.		

2. Johnson WL (2011). The validity and utility of the California Family Risk Assessment under practice conditions in the field: A prospective study. Child Abuse and Neglect 35: 18–28

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: Study aims to analyse the validity and implementation of a 'child maltreatment actuarial risk assessment model, the California Family Risk Assessment (CFRA)' (p18). Data extraction focuses on information about predictive validity - conceptualised as evaluation of prognosis within this study. **Description of theoretical approach?** Yes. Study uses concepts from medical prognostic models to examine validity of CFRA tool. **How was selection bias minimised?** No comparison group. Comparison group comprises same individuals as experimental group - but assessed using an additional instrument. **Was the allocation method followed?** Yes. **Is blinding an issue in this study?** No blinding. Study notes that workers were not blind to outcomes of earlier risk assessments, which may have led to confirmation bias.	**Were outcomes relevant?** Yes. **Were outcome measures reliable?** Yes. Substantiated maltreatment within two year of index incident. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Yes. Two-year follow-up. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. Comparison group comprises same individuals as experimental group - but assessed using an additional instrument. **Was intention to treat (ITT) analysis conducted?** Not reported. **Was the study sufficiently powered to detect an intervention**	**Does the study's research question match the review question?** Yes. Study relates to the validity and utility of a risk assessment tool. **Has the study dealt appropriately with any ethical concerns?** Partly. Study states that 'no human subjects review' of the project was required (assume this means ethics review). However, study used existing case records. Not clear whether participants were asked for consent to use their records for the purpose of this study. **Were service users involved in the study?** No. Service user records were used, but service users not involved in design or interpretation. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population is families reported for maltreatment.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity rating:** - Comparative information is the result of a 'natural experiment' occurring when practitioners choose to override the result of the CFRA, rather than a systematic comparison of practitioner judgements and CFRA. This is a relatively weak study design: ideally cases should have been assigned ratings using CFRA or practitioner judgements by 2 different individuals. Also, potential influence on risk of follow-up intervention is reported to have been statistically controlled for using logistic regression. However, the numbers of families receiving or not receiving intervention is not reported, making it difficult to judge whether this statistical adjustment is valid.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Did participants reflect target group? Yes. **Were all participants accounted for at study conclusion?** Yes.	**effect (if one exists)?** Not reported. Sample size for comparative element of study is relatively small (n=114) compared to main study sample (n=6,543). **Were the estimates of effect size given or calculable?** Yes. Odds ratios for each form of assessment predicted separately, however no direct comparison of the 2 forms of assessment. **Were the analytical methods appropriate?** Partly. CFRA and 'clinical' prediction compared through two logistic regression models. However, no direct statistical comparison of the 2 predictions, e.g. through entering both in to the same model. Furthermore, model aims to account for effects of intervention, however no data given on number of families receiving intervention so difficult to assess the validity of the model. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. Confidence intervals reported. **Do conclusions match findings?**	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Relates to Q7 assessment tools. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Partly. This study focuses specifically on predictive validity. **Does the study have a UK perspective?** No. US study (California).	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Conclusions are based on predictive validity of the model, and lack of predictive validity of worker 'override' judgements. However, analysis does not appear to directly compare them.		

Review question 7 – Findings tables

1. **Baumann DJ, Law JR, Sheets J et al. (2005). Evaluating the effectiveness of actuarial risk assessment models. Children and Youth Services Review 27(5): 465–90**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: A series of 3 field studies to examine the scientific integrity and practical utility of actuarial models of risk assessment in child welfare. Only data relating to studies 1 and 2 are reported here, as these looked at predictive validity, cross-validated using an objective measure. The third study did not look at examine predictive validity, and so was excluded. Study 1: RCT examining whether use of, or exposure to, an actuarial	**Participants:** Professionals/practitioners. Study 1: intake caseworkers; Study 2: investigation caseworkers. **Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Cases reports on physical abuse, sexual abuse, neglectful supervision and physical/medical neglect.	**Effect sizes:** Study 1: Intake decisions (outcome measures - correlation with case substantiation/confirmation). Correlation between risk caseworker judgements in the different conditions (computer assisted, new form and control) and substantiated abuse. Full data are not reported in the study. 1. Computer projections and confirmation (of abuse) at investigation showed a Spearman correlation of r=0.44 for physical abuse and r=0.24 for physical and medical neglect ($p<0.05$), suggesting that the actuarial model had some predictive validity. 2. Criterion validity of experimental group's judgements vs criterion validity of computer projections (only graphic data available [Fig 1, p473]): a. The correlations for substantiation for each type of maltreatment almost identical for the model and the computer group, strongly suggesting that exposure to the computer projections influenced the validity of the Com-	**Overall assessment of internal validity** - **Overall assessment of external validity** + **Overall validity score** - Reporting of analysis extremely unclear, including whether reported correlations are within-groups correlations between prediction and outcomes, or biserial correlations showing between-groups differences in

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
risk model led to superior judgements relating to substantiation of allegations compared to individuals not using or exposed to the model. Study 2: RCT examining whether use of, or exposure to, an actuarial risk model led to superior judgements relating to re-investigation compared to individuals not using or exposed to the model. **Methodology:** Study 1: RCT Study 2: RCT. **Country:** Not UK - USA and Canada. **Source of funding:** Government - USDHHS (US Department of Health & Human Services), ACF (The Administration for Children and Families).	• Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** Study 1: Total sample were 35 intake units (102 caseworkers and 2,625 maltreatment reports) randomised in 3 groups. Number of units randomised to control/intervention groups not reported. 1. Control group n reported as 180, however unclear what this number refers to and does not seem to tally with number of units, individuals or reports. 2. 'New form' group, n reported as 85, however unclear what this number refers to and does not seem to tally with number of units, individuals or reports 3. 'Computer group, n reported as 85, however unclear what this number refers to and does not seem to tally with number of units, individuals or reports. Study 2: Unit selection same as Study 1, involving 141 investigation caseworkers (discrepancies: as text also reported 135 from one region and 39 from another region, totalling 174 investigation caseworkers) and 979 families. No. of units randomised to control/intervention groups not reported. 1. Control group, n reported as 165, however unclear what this	puter group's judgements in the direction of the actuarial projections. b. The New Form group (not exposed to the actuarial projections) showed better validity than both the Control group and the computer group (only graphic data presented). The authors report that a statistical test of the difference between the New Form group and the Computer group, using Fisher's r to z transformations showed that New Form group showed better predictive validity (r=1.58, p<0.10 [borderline sig] for judgement on neglectful supervision; r=1.96, p<0.05 [sig] for judgment involving sexual abuse). However, it is unclear what the reported values represent - they are reported as 'r', but this test cannot take values outside +/-1. It is also unclear whether these represent within- or between-groups correlations. Study 2: Investigation decisions (outcome measures - correlation with case re-investigation). Full data are not reported in the study. Study reports that the investigation models are not as reliable as the intake model. 1. The authors state that the only model significantly correlated with actual re-investigation is physical neglect (r=0.12, p<0.07). However, it is unclear what criterion is being used. No data are reported for the other forms of abuse. 2. Case worker judgements in the New Form group were significantly correlated with re-investigation for all types of abuse (no statistical data reported).	the accuracy of prediction. Unclear what data points were used for analysis (cases, workers or units) and reported values of n for each condition do not appear to correspond to any of these.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	number refers to and does not seem to tally with number of units, individuals or reports. 2. New form group, n reported as 83, however unclear what this number refers to and does not seem to tally with number of units, individuals or reports. 3. Computer group, n reported as 215, however unclear what this number refers to and does not seem to tally with number of units, individuals or reports. Intervention numbers - see 'comparison number'. **Assessment tool** The risk assessment model had been developed in a prior study through statistical determination of features of cases which predicted substantiation or re-investigation. Study 1: Intake decisions (outcome measures - correlation with case substantiation/confirmation). 1. Control - works as usual, filled out paper copies of intake report (use of checklist of consensus-based items to determine case priority and written narrative. All dependent measures completed before formal decision, i.e. no chance to be influenced by actuarial models). 2. New Form group - works as usual plus also completed a paper actuarial risk form (items on this form had been		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	found to predict substantiation upon investigation). This group did not receive the actuarial feedback indicating the likelihood of substantiation provided to the Computer group. 3. Computer group - works as usual plus completing the computerised version of the actuarial risk form and the dependent measure prior to formal decision. Computer projections were used to inform these caseworkers about the likelihood of maltreatment substantiation based on their answers to the questions documented on the actuarial risk form. The four levels of risks were labelled 'Very Unlikely', 'Unlikely', 'Likely', and 'Very Likely' to result in substantiation of each of four different types of child maltreatment. Study 2: Investigation decisions (outcome measures - correlation with re-investigation). 1. Control - works as usual, filled out standard investigation report including the departmental risk assessment instrument. 2. New Form group - works as usual plus also completed a paper actuarial risk form (items on this form found to be related to future investigations within 14 months) and the departmental measures. All dependent		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	measures were related to investigators' estimate of the likelihood that the family would have another investigation within the next 9 months using the same four point scale 'Very Unlikely', 'Unlikely', 'Likely', and 'Very Likely' to result in substantiation of each of 4 different types of child maltreatment categories: physical abuse, sexual abuse, neglectful supervision, physical neglect. 3. Computer group - Same as 'New Form' Group, work as usual plus also completed an electronic version of the actuarial risk form. **Outcomes measured:** Study 1: correlation with substantiation/confirmation. Study 2: correlation with re-investigation.		

2. Johnson WL (2011). The validity and utility of the California Family Risk Assessment under practice conditions in the field: A prospective study. Child Abuse and Neglect 35: 18–28

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: Study aims to analyse the validity and implementation of a 'child maltreatment actuarial risk assessment model, the California Family Risk Assessment (CFRA)' (p18). Data extraction	**Participants:** Children and young people. Study utilises case records of families referred to California Child Welfare Services. Caregivers and families - Study utilises case records of families referred to California Child Welfare Services. **Sample characteristics:**	**Assessment** **Effect sizes** A - Logistic regression using CFRA risk scores Independent variables: CFRA risk score (reference category=low), post-investigation services provided (yes/no) Dependent variable: Substantiated maltreatment within 2 years of index incident 1. CFRA low risk - Wald statistic=7.94, p=0.02, no odds ratio (reference	**Overall assessment of internal validity** - **Overall assessment of external validity** + **Overall validity score**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
focuses on information about predictive validity - conceptualised as evaluation of prognosis within this study. **Methodology:** Overall design is a prospective evaluation. Comparative design is a smaller part of overall study. NCCSC data extraction has focused on this element. **Country:** Not UK, USA (California). **Source of funding:** Not reported.	• Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Some included cases may have had foster care as an intervention - not clear from reporting. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** Intervention numbers - 114 cases were assessed using the California Family Risk Assessment and subsequently overridden using clinical judgement. **Assessment tool:** California Family Risk Assessment is an actuarial risk assessment model developed by the Children's Research Centre. CFRA is completed by Child Welfare Workers following a 3-hour training course. CFRA comprises 2 10-item scales: 1 assesses future likelihood of future	category) 2. CFRA moderate risk - Wald statistic=0.04, p=0.85, OR=1.2 (95% CI 0.22 to 6.50) 3. CFRA high risk - Wald statistic=4.48, p=0.03, OR=6.3 (95% CI 1.15 to 34.78) 4. Post-investigation services provided - Wald statistic=0.37, p=0.55, OR=1.6 (95% CI 0.37 to 6.56) 5. Constant - Wald statistic=9.51, p=0.002, OR=0.051 (95% CI not reported) B - Logistic regression using clinical judgement 'override' risk scores Independent variables: Clinical judgement 'override' risk score (reference category=low/moderate), post-investigation services provided (yes/no) Dependent variable: Substantiated maltreatment within 2 years of index incident 1. Clinical judgement risk score low/moderate - Wald statistic=0.5, p=0.98, no odds ratio (reference category) 2. Clinical judgement risk score high -Wald statistic=0.02, p=0.88, OR=1.16 (95% CI 0.16 to 8.28) 3. Clinical judgement risk score very high - Wald statistic =0.05, p=0.82, OR=1.21 (95% CI 0.24 to 6.23) 4. Post-investigation services provided - Wald statistic =0.03, p=0.86, OR=1.13 (95% CI 0.28 to 4.63) 5. Constant - Wald statistic=6.57, p=0.01, OR=0.11 (95% CI not reported)	- Comparative information is the result of a 'natural experiment' occurring when practitioners choose to override the result of the CFRA, rather than a systematic comparison of practitioner judgements and CFRA. This is a relatively weak study design: ideally cases should have been assigned ratings using CFRA or practitioner judgements by two different individuals. Also, potential influence on risk of follow-up intervention is reported to have been statistically controlled for using logistic regression. However, the numbers of families receiving or not receiving intervention is not reported, making it difficult to judge whether this statistical adjustment is valid.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	physical or sexual abuse, 1 assesses future likelihood of neglect. Scales result in a score of low-, moderate-, high- or very high risk. The highest score on either scale forms the basis for decisions about what services are provided. The results of the CFRA are used to decide whether to provide 'in-home' child protection services and the intensity of support provided. **Comparison tool/usual practice:** The comparator in this study were clinical judgements on level of risk, made after workers decided to 'override' the CFRA rating. Worker judgements can take the following forms: 1) a one-category increase when the worker's impressions suggest that the case is higher risk than CFRA indicates and 2) changing the category to 'very high risk' in the presence of particular indicators (indicators not reported).		

Review question 8 – What aspects of professional practice support and hinder effective assessment of risk and need in relation to child abuse and neglect?

Review question 8 – Critical appraisal tables

1. **Brandon M, Belderson P, Warren C et al. (2008) Analysing child deaths and serious injury through abuse and neglect: what can we learn? - A biennial analysis of serious case reviews 2003–2005. London: Department of Education**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
1. Qualitative component 1 Which component? Thematic analysis of 47 SCR reports. **1.1 Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Partly. Reason for selecting these 47 for further analysis appears to be convenience sampling - these were the reports for which the full Overview Reports were available. **1.2 Is the process for analysing qualitative data relevant to address the research question?** Partly. Unclear how 'emerging themes' were identified and verified. **1.3 Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. Little consideration of	**4. Quantitative component description A (including incidence or prevalence study without comparison group; case series or case report)** **Which component?** Collection and analysis of data from total of 161 case reviews. **4.1. Is the sampling strategy relevant to address the quantitative research question (quantitative aspect of the mixed-methods question)?** Yes. Sample comprises all 161 SCRs published between 2003 and 2005. **4.2 Is the sample representative of the population under study?** Yes. **4.3. Are measurements appropriate (clear origin, or validity known, or standard instrument)?** Yes. Key characteristics of children and families who are the subjects of the SCRs.	**Does the study's research question match the review question?** Partly. The study has a series of objectives, one of which is to 'identify any lessons for policy and practice, including examples of good practice' - this is considered to be relevant to Q14. **Has the study dealt appropriately with any ethical concerns?** No. No mention of ethical approval process, although potentially of lower concern as secondary analysis of documentary sources, rather than primary research with service users. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population is Serious Case	**Overall assessment of internal validity** + **Overall assessment of external validity** ++ **Overall validity score** + Key limitations of the study are a lack of clarity with respect to the way in which thematic analysis of the sub-sample of 47 reviews was conducted, and how the findings from this analysis has been integrated with quantitative analysis (see Chapter 6). However, study strength is that there is a 100% sample of SCRs from the 2003–5 time period.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
how themes are linked or otherwise to other elements of the cases. **1.4 Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?** N/A. Documentary analysis.	**4.4. Is there an acceptable response rate (60% or above)?** Yes. **5.1. Is the mixed-methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and quantitative aspects of the mixed-methods question?** Yes. **5.2. Is the integration of qualitative and quantitative data (or results) relevant to address the research question?** Partly. There is some synthesis of qualitative and quantitative components, for example in Chapter 6. However, the process by which the data were integrated is not clear. **5.3 Is appropriate consideration given to the limitations associated with this integration, such as the divergence of qualitative and quantitative data (or results)?** No.	Reviews about cases in which children have experienced abuse and neglect (leading to death or significant harm). **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study has information about multiple aspects of practice, including early help. **g. Does the study have a UK perspective?** Yes.	

2. **Cleaver H and Walker S (2004) From policy to practice: the implementation of a new framework for social work assessments of children and families. Child & Family Social Work 9: 81–90**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. To explore views of social work practitioners. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Defensible. Postal questionnaires and interviews. **How well was the data collection carried out?** Not sure/inadequately reported. Postal questionnaires and phone interviews, and meetings with managers and staff. (Limited details on how these were designed and executed.)	**Is the context clearly described?** Unclear. Limited information on characteristics of participants. However, study conducted within a context of existing challenges due to organisational change, difficulties in staff retention and recruitment, and changes in and adaptation to information technology such as electronic data processing and recording. **Was the sampling carried out in an appropriate way?** Not sure. No information on how sample selected. **Were the methods reliable?** Not sure. Data collected by postal questionnaires and interviews, limited details. **Are the data 'rich'?** Rich. **Is the analysis reliable?** Somewhat reliable. Descriptive and narrative analysis, no details. **Are the findings convincing?** Somewhat convincing. Limited methodological details. **Are the conclusions adequate?**	**Does the study's research question match the review question?** Yes. To evaluate the implementation of the Framework for the Assessment of Children in Need and their Families. To explore impact of the Assessment Framework on practice (the focus of this review). **Has the study dealt appropriately with any ethical concerns?** No. Not reported. **Were service users involved in the study?** No **Is there a clear focus on the guideline topic?** Yes. Q8: To explore impact of the Assessment Framework on practice. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Social work practitioners, managers, and staff from partner agency; also parents and children. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall validity rating:** - Limited methodological details on postal questionnaires (what questions asked? response rates) and interviews (how and where conducted, what questions asked, specifics of participants and interviewer etc.). The author stated that 'Because the authors were responsible for some of the research that informed the Assessment Framework and were involved in the development of the assessment records, this study does not purport to be an objective evaluation' (p83).

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Somewhat adequate. Limited methodological details	**Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Assessment of child maltreatment. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. Views from social work practitioners/professionals in child protection. **Does the study have a UK perspective?** Yes. The Framework for the Assessment of Children in Need and their Families, in the UK.	

3. **Devaney J, Bunting L, Hayes D et al. (2013) Translating Learning into Action: An overview of learning arising from Case Management Reviews in Northern Ireland 2003-2008. Belfast: Queen's University Belfast**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Thematic analysis of Case Management Reviews. **Is the study clear in what it seeks to do?** Mixed. The study seeks to identify 'key themes' across the 24 Case Management Reviews. It is not specified what nature of issues could be considered within this category.	**Is the context clearly described?** Clear. Contextual information about cases reviewed is provided. **Was the sampling carried out in an appropriate way?** Appropriate - 100% sample of all CMRs in a given time period. **Were the methods reliable?** Reliable. Learning from the reviews	**Does the study's research question match the review question?** Partly. The study's research question is about identifying 'key learning' from Case Management Reviews. Part of this involves thematic analysis of 'key themes' which include issues relevant to aspects of professional practice.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
How defensible/rigorous is the research design/methodology? Somewhat defensible. Limited justification of analytic techniques provided. **How well was the data collection carried out?** Appropriately. Analysis of CMR reports.	has been triangulated with relevant research evidence as appropriate. **Are the data 'rich'?** Mixed. There are relatively few examples given to illustrate the themes identified. **Is the analysis reliable?** Somewhat reliable. It is unclear how the thematic analysis was undertaken, therefore difficult to judge reliability of analysis. **Are the findings convincing?** Somewhat convincing. Themes identified are often supported by other aspects of research literature. **Are the conclusions adequate?** Adequate.	**Has the study dealt appropriately with any ethical concerns?** Yes. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Partly. Consideration of professional practice and ways of working forms part of the analysis conducted in the study. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population includes children and young people who, at one point, showed early signs of abuse and neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. **Does the study have a UK perspective?** Yes	Overall, there is a lack of description of how thematic analysis was undertaken.

4. **Horwath J (2005) Identifying and assessing cases of child neglect: learning from the Irish experience. Child and Family Social Work 10: 99–110**

Internal validity - approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear. Analysis of case files and views and experiences of social workers. **How defensible/rigorous is the research design/methodology?** Defensible. Analysis of case files and views and experiences of social workers. **How well was the data collection carried out?** Appropriately. Data on assessment process from case reviews; Vs & Es from surveys and focus groups.	**Is the context clearly described?** Not sure. Study conducted over 10 years ago, practice and views may have changed. **Was the sampling carried out in an appropriate way?** Appropriate. Social workers from teams covering both large towns and remote farming communities. **Were the methods reliable?** Somewhat reliable. For case reviews: piloting and content analysis to develop coding frame for recording data. For social workers' responses: focus groups and questionnaires (limited details on these methods). **Are the data 'rich'?** Rich. **Is the analysis reliable?** Reliable. Questions within the questionnaire and the focus groups were used to validate and clarify findings from the case audit/reviews. **Are the findings convincing?** Convincing. **Are the conclusions adequate?**	**Does the study's research question match the review question?** Yes. Professional practice support and hinder relating to effective assessment of risk and need in child maltreatment. **Has the study dealt appropriately with any ethical concerns?** No. Not reported. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. Professional practice in relation to assessment. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Social workers, including managers. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by**	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + Limited methodological details of questionnaire surveys and focus groups. Study was conducted over 10 years ago and practice would have changed since.

Internal validity - approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Adequate.	**the guideline?** Yes. Risk assessment process. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. Analysis of case files and views and experiences of social workers. **Does the study have a UK perspective?** No. Ireland.	

5. **London Safeguarding Children Board (2011) Final monitoring report: local authority pilots of the London safeguarding trafficked children guidance and toolkit. London: London Safeguarding Children Board**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
1. Qualitative component 1 Which component? Qualitative survey. **1.1 Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. **1.2 Is the process for analysing qualitative data relevant to address the research question?** No. No information given regarding how qualitative data have been analysed. Also unclear who completed survey, what their job role was, and using what process	**3. Quantitative component (incl. non-RCT; cohort study; case-control study)** Which quantitative component? Two quantitative surveys. **3.1 Are participants (organisations) recruited in a way that minimises selection bias?** Unclear. Process for recruiting pilot local authorities, and survey respondents within those authorities, is not reported. **3.2 Are measurements appropriate (clear origin, or validity known, or standard instrument; and absence of contamination**	**Does the study's research question match the review question?** Partly. Study is about monitoring use of guidance and a toolkit for identifying and assessing children who have been trafficked. Includes some information about practitioner views of the toolkit, and what helps and hinders assessment of trafficked children. **Has the study dealt appropriately with any ethical concerns?** Partly. No ethical approval sought. However, study did not directly involve service users.	**Overall assessment of internal validity** - **Overall assessment of external validity** + **Overall validity score** - Very poor information regarding methodology, including how pilot sites were recruited, which individuals participated within those sites, and how results were analysed. Incomplete reporting of results and lack of clarity regarding

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
(i.e. did they consult with other members of staff). **1.3 Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. **1.4 Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?** No. **1. Qualitative component 2** Which component? National monitoring workshop with 10 of the 12 pilot authorities. **1.1 Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. **1.2 Is the process for analysing qualitative data relevant to address the research question?** No. No information given regarding how data from the workshop was recorded or analysed. **1.3 Is appropriate consideration given to how findings relate to**	**between groups when appropriate) regarding the exposure/intervention and outcomes?** No. **3.3 In the groups being compared (exposed versus non-exposed; with intervention versus without; cases versus controls), are the participants comparable, or do researchers take into account (control for) the difference between these groups?** N/A. **3.4 Are there complete outcome data (80% or above), and, when applicable, an acceptable response rate (60% or above), or an acceptable follow-up rate for cohort studies (depending on the duration of follow-up)?** N/A. **5.1. Is the mixed-methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and quantitative aspects of the mixed-methods question?** Yes. **5.2. Is the integration of qualitative and quantitative data (or results) relevant to address the research question?** No. Not always clear which data sources are	**Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population is practitioners working with children and young people at risk of or experiencing abuse and neglect (trafficking). **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to recognition and assessment. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes.	elements of data collection have contributed to which finding.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
the context, such as the setting, in which the data were collected? No. **1.4 Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?** No. **1. Qualitative component 3** Which component? Multi-agency workshop with practitioners. However, not clear how the findings of this were used. **1.1 Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. **1.2 Is the process for analysing qualitative data relevant to address the research question?** No. Unclear how participants were recruited, how data recorded and how analysed. **1.3 Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No.	being referred to for particular findings. **5.3 Is appropriate consideration given to the limitations associated with this integration, such as the divergence of qualitative and quantitative data (or results)?** No.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
1.4 Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants? No.			

6. Ofsted (2014) In the child's time: professional responses to neglect. Manchester: Ofsted

Internal validity - approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Mixed. The report explores the effectiveness of arrangements to safeguard children who experience neglect. A mixture of designs involved in a thematic inspection by Ofsted exploring the response of professionals when they identify neglect: case reviews, auditing, survey, interviews etc. (This review will focus on the assessment aspects to highlight what support/hinder effective assessment of maltreatment risk in children.) **How defensible/rigorous is the research design/methodology?** Defensible. Interviews /discussions with users.	**Is the context clearly described?** Clear. **Was the sampling carried out in an appropriate way?** Not sure. Not reported. **Were the methods reliable?** Not sure. Limited details provided. **Are the data 'rich'?** Mixed. **Is the analysis reliable?** Not sure/not reported. **Are the findings convincing?** Somewhat convincing. **Are the conclusions adequate?** Somewhat adequate.	**Does the study's research question match the review question?** Yes. A survey and case review of practice, draws on evidence from cases, also views of parents, carers and professionals from the local authority and partner agencies. **Has the study dealt appropriately with any ethical concerns?** No. **Were service users involved in the study?** Yes. Services users (parents, carers) participated in this report. **Is there a clear focus on the guideline topic?** Yes. Views of parents, carers and professionals from the local authority and partner agencies. **Is the study population the same as at least 1 of the groups**	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity rating:** - Limited reporting of data collection methods.

Internal validity - approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
How well was the data collection carried out? Not sure/inadequately reported.		**covered by the guideline?** Yes. Parents, carers and professionals. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Explores the effectiveness of arrangements to safeguard children who experience neglect. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes.	

7. **Platt D (2008) Care or control? The effects of investigations and initial assessments on the social worker-parent relationship. Journal of Social Work Practice 22: 301–15**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Study seeks to understand in-depth experiences of assessment, therefore qualitative approach appropriate (Patton 1990). The 23 qualitative interviews enrich the perspective of both social worker and parent in the child protection experience.	**Is the context clearly described?** Unclear. The author provides explicit information about the characteristics of the participants: 'the sample included a high proportion of lone mothers, and a range of children's ages from pre-birth to 16 ... The characteristics were	**Does the study's research question match the review question?** Yes. The study is in relation to the impact of assessments, i.e. Initial (S.17) and investigation i.e. (S.47) has on social worker and parent relationships.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall validity rating:**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Under each finding are quotations from a parent and social worker to illuminate key aspects and lived experiences of where there has been effective and ineffective assessments and why this might be. **Is the study clear in what it seeks to do?** Clear. Clear research question relation to the effects of investigations and initial assessments on social worker-parent relationships. The study has clear objectives which is to explore assessment using the following terminology: coercive (S.47 investigations) and less coercive (S.17 initial assessment), and poses whether the formalised assessment process is the reason for poor relationships between social worker-parent. The study concludes by asserting that the role of skilled workers are at least 'equally important' (p314). The literature is appropriate to set the current context of a shift from investigations to the use of initial assessments, and comparisons are made with the therapeutic alliance in the counselling and psychotherapy literature, to highlight less coercive relationships.	reasonably consistent with similar research (Tunstill & Aldgate 2000; Cleaver & Walker, 2004)' (p305). However, there is less consideration of the impact of the contexts of the two case study sites. This is potentially highly relevant, given that there may have been idiosyncrasies in assessment practices in the 2 sites. **Was the sampling carried out in an appropriate way?** Appropriate. The sample is purposive to incorporate both more coercive and less coercive assessment processes. It is clear that the authors were identifying cases that were involving families where there were concerns about the child protection threshold. **Were the methods reliable?** Somewhat reliable. The data was collected by one method, which were qualitative interviews, however there was not further comparisons made amongst other studies. **Are the data 'rich'?** Rich. The data is generally well founded in grounded theory and analysed in NVivo. The authors give justifica-	**Has the study dealt appropriately with any ethical concerns?** Yes. The Research Committee of the Association of Directors of Social Services granted ethical approval. Parents consented to be interviewed and 2 refused to be. Alternative neutral locations were provided for parents if they did not want a home visit. **Were service users involved in the study?** No. Service users have not co-researched this paper. **Is there a clear focus on the guideline topic?** Yes. The study has a focus in relation to the guideline topic because the findings explore what supports and hinders assessment, particularly on the impact of the relationship between social worker and parent. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. The population of the study is qualitative interviews with parents whose children are undergoing child protection assessments/investigations (at risk of/or experiencing abuse and neglect). The study also interviews practitioners	- Overall, the study meets most of the quality criteria. However, little consideration of impact of the working context in the 2 case study local authorities. This is potentially highly relevant, given that there may have been idiosyncrasies in assessment practices in the two sites. Relatively small number of investigations (n=3) compared to initial assessments (n=20) on which to form a basis of comparisons.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
How defensible/rigorous is the research design/methodology? Somewhat defensible. Purposive sampling to explore cases in which concerns came close to child protection thresholds, and to include both initial assessments and investigations. However, relatively small number of cases involving investigation (n=3). The discussion explores the limitations of interviewing predominantly females, and the nature of assessment and sample group is limited to initial assessment and borderline child protection risk. **How well was the data collection carried out?** Not sure/inadequately reported. Little detail regarding data collection is provided, only that interviews were conducted using an interview guide approach. Authors are explicit that interviews were recorded subject to consent of interviewees.	tion with other research and explore why this approach helped highlight 'key elements of social worker-parent relationship'. In addition, these elements are explored and comparisons made throughout between more coercive and less coercive interventions, thus meeting the research aim. **Is the analysis reliable?** Not sure/not reported. Process of analysis is not reported further than using grounded theory and NVivo software. **Are the findings convincing?** Convincing. The findings are clearly presented and are supported by extracts from data collection. Comparisons are made and author meets the research aim in a coherent way illuminating three key aspects of positive social worker-parent relationship during more coercive and less coercive interventions: 'sensitivity, honesty and straightforwardness, and listening and accurate understanding' (p306). **Are the conclusions adequate?** Adequate. The conclusions are well founded in the context of the	who have worked with these families at the point of referral. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Relates to assessment. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. The aim explores 'relationship aspects of the study- between social workers and parents - and considers the question of whether improvements may be achieved in these relationships if less coercive, more supportive responses are used' (p305). **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. The findings explore social workers approach to what supports and hinders effective assessment of risk and need of children through interviews with both social workers and families. The study highlights three key elements of	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	paper, and a discussion of the limitation of the research in terms of generalisability is alluded to. The authors postulate that the variation in more coercive and less coercive intervention/assessment was not a 'prerequisite for good relationships, skilled workers developed good working relationships in a variety of procedural contexts' (p.313).	the social worker-parent relationship to be theorised as 'sensitivity, honesty and straightforwardness, and listening and accurate understanding' (p306). **Does the study have a UK perspective?** Yes. The study setting is 2 urban North England social services. However, caution to generalise the whole of UK.	

8. **Rigby P (2011) Separated and trafficked children: The challenges for child protection professionals. Child Abuse Review 20: 324–40**

Internal validity – approach and sample.	Internal validity – performance and analysis.	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Study seeks to understand challenges in practice. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Defensible. **How well was the data collection carried out?** Somewhat appropriately. Lack of clarity regarding interviewees compared to focus groups, and which topics were discussed via which method.	**Is the context clearly described?** Unclear. Little description of context of child trafficking in Glasgow, for example the extent of the issue, whether this is a relatively new phenomenon etc. **Was the sampling carried out in an appropriate way?** Not sure. Method for selecting case files and interviewees not reported. **Were the methods reliable?** Reliable. **Are the data 'rich'?** Not sure. Little presentation of data within the report.	**Does the study's research question match the review question?** Partly. Main research question is about identifying challenges emerging for practitioners working with separated children who have been trafficked. However, there is a section on 'Identification and assessment'. **Has the study dealt appropriately with any ethical concerns?** Partly. Author states that research governance provided by local Child Protection Committee but does not appear that any formal research ethics approval sought. Does not appear that consent was	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + Assessment is a subset of overall research question. **Overall validity rating:** Not clear which data were gathered via interview, and which via focus group. Analysis methods unclear. Relatively little reference to, or presentation of, primary data gathered.

Internal validity – approach and sample.	Internal validity – performance and analysis.	External validity	Overall validity rating
	Is the analysis reliable? Somewhat reliable. Analysis method described as 'broadly grounded theory'. Analysis does not appear to have been cross-validated by a second researcher or similar. **Are the findings convincing?** Somewhat convincing. Little presentation of primary data to back up points made in analysis. **Are the conclusions adequate?** Adequate.	obtained from children for using their case files for analysis purposes. **Were service users involved in the study?** No. Research not co-produced by service users. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least one of the groups covered by the guideline?** Yes. Population is children who have been trafficked and the professionals working with them. **Is the study setting the same as at least one of the settings covered by the guideline?** Yes. **Does the study relate to at least one of the activities covered by the guideline?** Yes. Study relates to assessment. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes. Scotland (Glasgow).	

9. **Robertson A (2014) Child welfare assessment practices in Scotland: an ecological process grounded in relationship-building. Journal of Public Child Welfare 8: 164–89**

Internal validity – approach and sample.	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. Clear theoretical rationale for mixed methods case study approach. However, in practice relatively small sample size and little consideration of how results may vary in the 2 case study areas. **How well was the data collection carried out?** Appropriately.	**Is the context clearly described?** Clear. **Was the sampling carried out in an appropriate way?** Not sure. Unclear why particular individuals included in sample - no overall sampling frame. **Were the methods reliable?** Somewhat reliable. Individual data collection methods appear reliable, but process of 'triangulating' data not clearly described. **Are the data 'rich'?** Mixed. Little exploration of impact of different case study sites. **Is the analysis reliable?** Unreliable. Unclear how thematic analysis of semi-structured interviews has been conducted to arrive at four themes of relationship characteristics that support the assessment process. No description of analysis process, whether any qualitative analysis software was used. **Are the findings convincing?** Somewhat convincing.	**Does the study's research question match the review question?** Partly. Study has 3 research questions, 1 of which is relevant to our review question: 1. What are Scottish child welfare experts' views of key relationship characteristics during the assessment process? Two questions not relevant: 1. How do new, Scottish child welfare policies and specifically GIRFEC address assessment and child wellbeing? 2. What are Scottish child welfare experts' practice experiences with Getting It Right For Every Child (GIRFEC). **Has the study dealt appropriately with any ethical concerns?** Yes. Institutional Review Board approval obtained, and informed consent from participants. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population is professionals	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall validity rating:** - Unclear how thematic analysis of semi-structured interviews has been conducted to arrive at 4 themes of relationship characteristics that support the assessment process. No description of analysis process, whether any qualitative analysis software was used.

Internal validity – approach and sample.	Internal validity – performance and analysis	External validity	Overall validity rating
	Are the conclusions adequate? Somewhat adequate.	working with children and young people at risk of, or experiencing, abuse and neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to assessment. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes. Scotland.	

10. **Selbie J (2009) Health visitors' child protection work: exploratory study of risk assessment. Community practitioner: the journal of the Community Practitioners' and Health Visitors' Association 82(5), 28–31**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. To elicit opinions from focus groups. **Is the study clear in what it seeks to do?** Clear. To identify facilitators and enablers in identifying risks to children.	**Is the context clearly described?** Clear **Was the sampling carried out in an appropriate way?** Not sure. Not reported. **Were the methods reliable?** Reliable. Semi-structured interviews with open-ended questions.	**Does the study's research question match the review question?** Yes. To seek health visitors' (HV) opinions on the efficacy of health assessment and screening tools in child protection work. **Has the study dealt appropriately with any ethical concerns?**	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
How defensible/rigorous is the research design/methodology? Defensible. **How well was the data collection carried out?** Appropriately. 2 focus groups and 1:1 interview.	**Are the data 'rich'?** Rich. **Is the analysis reliable?** Somewhat reliable. Thematic analysis using grounded theory approach. Interview data audio-taped and transcribed; context noted and fieldnotes collected. **Are the findings convincing?** Somewhat convincing. **Are the conclusions adequate?** Somewhat adequate.	Yes. Ethical approval sought from NHS Research Ethics Committee. **Were service users involved in the study?** Yes. Health professionals (HV) participated in study. **Is there a clear focus on the guideline topic?** Yes. Views and opinions of HV in child protection work. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Professionals - health visitors. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Not applicable (views study). **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Assessment/screening tools. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. From HV on efficacy of tools. **Does the study have a UK perspective?** Yes. UK.	

11. Sen R, Lister PG, Rigby P et al. (2014) Grading the Graded Care Profile. Child Abuse Review 23: 361–73

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
1. Qualitative component 1 Which component? Focus groups with practitioners. **1.1 Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. **1.2 Is the process for analysing qualitative data relevant to address the research question?** Partly. Limited description of analysis 'focus groups were ... coded according to emergent themes, which were subsequently refined until they fully encapsulated the interview dataset' (p365). **1.3 Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. Very little data regarding focus groups. **1.4 Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?**	**3. Quantitative component (incl. non-RCT; cohort study; case-control study)** Which quantitative component? Survey. **3.1 Are participants (organisations) recruited in a way that minimises selection bias?** Unclear. Not reported. **3.2 Are measurements appropriate (clear origin, or validity known, or standard instrument; and absence of contamination between groups when appropriate) regarding the exposure/intervention and outcomes?** Unclear. No information on questionnaire items. **3.3 In the groups being compared (exposed versus non-exposed; with intervention versus without; cases versus controls), are the participants comparable, or do researchers take into account (control for) the difference between these groups?** N/A. **3.4 Are there complete outcome data (80% or above), and, when applicable, an acceptable response rate (60% or above), or**	**Does the study's research question match the review question?** Yes. Study relates to practitioner views on an assessment tool. **Has the study dealt appropriately with any ethical concerns?** Partly. Research was approved by the University of Strathclyde Ethics Committee and the authority's research committee. Parent participants gave consent. Unclear whether practitioners were asked for informed consent. **Were service users involved in the study?** No. Service users involved as participants, but did not co-produce research. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least one of the groups covered by the guideline?** Yes. Parents of children at risk of or experiencing neglect. Professionals working with children at risk of or experiencing neglect.	**Overall assessment of internal validity** - **Overall assessment of external validity** ++ **Overall validity score** - Very little information given regarding data collection or analysis for any of the data collection methods. It is not always clear what data source findings are based on.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
No. No consideration of researcher influence. **1. Qualitative component 2** Which component? Semi-structured interviews with practitioners who had cases where GCP should have been used (n=56). **1.1 Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. **1.2 Is the process for analysing qualitative data relevant to address the research question?** Partly. Limited information provided about analysis process. **1.3 Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. Not considered. **1.4 Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?** No. **1. Qualitative component 3**	**an acceptable follow-up rate for cohort studies (depending on the duration of follow-up)?** Unclear. Contacted sample not reported - only achieved sample. **5.1. Is the mixed-methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and quantitative aspects of the mixed-methods question?** Partly. Unclear what questionnaire data have added to understanding. **5.2. Is the integration of qualitative and quantitative data (or results) relevant to address the research question?** Partly. **5.3 Is appropriate consideration given to the limitations associated with this integration, such as the divergence of qualitative and quantitative data (or results)?** No.	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least one of the activities covered by the guideline?** Yes. Study relates to assessment. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes. Scotland.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Which component? Interviews with families where GCP had been used, observation of families and follow-up interviews. **1.1 Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. **1.2 Is the process for analysing qualitative data relevant to address the research question?** Partly. Limited information regarding data collection and analysis. **1.3 Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. **1.4 Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?** No.			

12. Vincent S and Petch A (2012) Audit and Analysis of Significant Case Reviews. Edinburgh: The Scottish Government

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Content analysis of SCR reports.	**Is the context clearly described?** Clear. Limitations of	**Does the study's research question match the review question?**	**Overall assessment of internal validity:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is the study clear in what it seeks to do? Clear. **How defensible/rigorous is the research design/methodology?** Defensible. **How well was the data collection carried out?** Appropriately. Relevant content extracted from reports using template.	SCRs as a source of data made clear. **Was the sampling carried out in an appropriate way?** Appropriate. All SCRs published in a particular timeframe (post 2007). **Were the methods reliable?** Somewhat reliable. Template used for analysing reports. However, no mention of double coding or cross-validation by a second member of the team. **Are the data 'rich'?** Mixed. Little contextualisation of findings in the context of cases. Some direct quotes from SCR reports used. **Is the analysis reliable?** Not sure/not reported. Little data presented on which to base this judgement. **Are the findings convincing?** Somewhat convincing. Little presentation of primary data to show how particular themes/issues have been identified. **Are the conclusions adequate?** Somewhat adequate.	Partly. Overall research question is about learning from Serious Case Reviews, but there is one section relating to assessment. **Has the study dealt appropriately with any ethical concerns?** Partly. Steps taken to ensure that information from SCR reports remained anonymised. No mention of ethical approval. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study relates to cases where children have died or been injured, the majority of which had an element of abuse or neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study as information relevant to assessment.	**Overall assessment of external validity:** + Information on assessment is part of a broader study. **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Does the study have a UK perspective? Yes. Scotland.	

Review question 8 – Findings tables

1. **Brandon M, Belderson P, Warren C et al. (2008) Analysing child deaths and serious injury through abuse and neglect: what can we learn? - A biennial analysis of serious case reviews 2003–2005. London: Department of Education**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The aims of the study are: 'i. To provide descriptive statistics from the agreed full sample (i.e. 161 cases), illustrated by some examples from the reviews ... ii. To scrutinise a sub sample if cases (i.e. 47) to chart thresholds of multi-agency intervention at the levels specific in Every Child Matters (Cm 5860 2003) ... iii. Building on the learning from the first two objectives, to seek a meaningful analysis by identifying some ecological-transactional factors within the sub-sample of reviews ... iv. To provide	**Participants** Children and young people. Sample comprises: 161 Serious Case Review reports, conducted 'when abuse and neglect are known or suspected factors when a child dies (or is seriously injured or harmed), and there are lessons to be learnt about inter-agency working to protect children' (p7). The 161 SCRs studied were notified during the period April 2003 to March 2005. **Sample characteristics** • Age - The ages of the children who were subject to SCRs considered in the study were as follows: 0–1 month - 13% 2–13 months - 19% 4–6 months - 11% 7–12 months - 4% 1–3 years - 18% 4–5 years - 2% 6–10 years - 7% 11–15 years - 16% 16 years + - 9% (n not given - assume 161).	**Narrative findings** Data extracted from Chapter 4 – 'Assessment and analysis: an ecological - transactional perspective'. In their preliminary discussion the authors highlight the importance of the Framework for the Assessment of Children in Need and their Families (Department of Health 2000), noting that this is based on an ecological model. They go on to emphasise the distinction between descriptive and dynamic analysis, a concept that is drawn on to distinguish between high quality and poor quality assessment practices examined in the serious case reviews. Of particular concern to the authors is the failure of practitioners to establish '… a thorough social history on which to base a more coherent and developmentally informed analysis (rather than description) …' (p56). The authors go on to suggest that in contrast to guidance set out in the Framework for the Assessment of Children in Need and their Families (Department of Health 2000), practitioners analysis of risk in the cases under examination tended to be static or non-dynamic. This in turn led to dense descriptions of facts rather than clear explanations. The report also	**Overall assessment of internal validity** + **Overall assessment of external validity** ++ **Overall validity score** + Key limitations of the study are a lack of clarity with respect to the way in which thematic analysis of the sub-sample of 47 reviews was conducted, and how the findings from this analysis has been integrated with quantitative analysis (see Chapter 6). However, study strength is that

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
practice tools for use by Local Safeguarding Children Boards and practitioners and to identify any lessons for policy and practice, including examples of good practice' (p15). **Methodology:** Other. Analysis of Serious Case Reviews - analogous to thematic analysis of multiple case studies - therefore use qualitative study critical appraisal tool. **Country:** UK. **Source of funding:** Government.	• Sex - Female - 55% Male - 45% (n not given - assume 161). • Ethnicity - White/White British (74%) Mixed (6%) Black/Black British (13%) Asian/Asian British (6%) Other ethnic group (1%) (n=136). • Religion/belief - Not reported. • Disability - Disability recorded in 5% of cases (n=161). • Long term health condition - Information on long term health conditions available for 'intensive' sample only (n=47) Complex health needs - 9% Chronic illness - 11%. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Head injury - 16% Sudden Infant Death < 4% Overlying - 4% Physical assault - 35% Neglect - 21% Poisoning/overdose - 4% Suicide - 9% Sexual abuse - 4% Gone missing - 4% Other <4% (n=161). • Looked after or adopted status - In care at time of incident - 10% (n=159). • Unaccompanied asylum seeking, refugee or trafficked children - Not reported.	emphasises the importance of '… detailed descriptions of the parent's developmental, attachment and relationship history …' (p63) which the authors note is often missing from case records examined by reviews. Similarly, they note that in many of these cases, assessments simply recorded what was happening in a case rather than analysing why these things were happening. The authors go on to comment that '… it was all too common for agencies and assessors to "describe" their way around the three sides of the Assessment Framework without properly generating an analysis or formulation of what was happening at the psychosocial level between the key actors, including the professionals themselves' (p56). The report provides an analysis of the quality and quantity of information collected in 42 cases of the 'intensive sample'. The authors note that in cases where there was a mother/female carer or father/male carer involved '… there was little, if any information about the carer's own developmental and relationship history. The absence of information about the parent's developmental and relationship history is likely to limit the value, usefulness and insightfulness of any assessment. Although descriptions of current parenting behaviours (capacities) are necessary, on their own they lack the dynamic quality achieved when a psychological and historical perspective is taken' (p56). "Inferred levels of information collected and available in each case and its analysis, assessment and formulation' (n=42) – 'quantity of information collected by or available to all agencies, whether shared or not' (p64). The authors also note that the cases illustrated the long-standing issue of poor information sharing	there is a 100% sample of SCRs from the 2003–5 time period.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Sample size - Main sample n=161. SCR reports Intensive sample n=47 SCR reports.	between agencies with 'full sharing' being identified in only 10% of cases. The report states that most of the reviews had judged practitioner's assessments and case analysis as weak. Reasons for this included a failure to explicitly undertake an assessment; the failure to collect sufficient levels of evidence and information; a failure to analyse the evidence clearly in order to develop an explanation for or understanding of what was happening in the case. The failure to explore the '… interactive and diagnostic effects of vulnerability and risk, resilience and protective factors …' is also highlighted; and many assessments are described as '… little more than the accumulation and presentation of disparate facts and information' (p 65). As one serious case review stated: 'In the plans for the children there was no indication that the issues emerging from [the mother's] past, and which clearly impacted upon her ability to parent, were addressed. Her issues of loss, the impact of her sexual abuse experiences on her ability to form relationships, the complex relationship with her mother should have featured in the assessment and decision making processes and have been the subject of specific detailed work with her. The fact that this did not occur is of major concern' (Case details unclear, quoted on p65). The report notes that consideration of the three domains of the Assessment Framework was identified as inadequate in 76% of cases in the intensive sample. This is deemed to be essential to a robust analysis of risk as it enables a holistic and ecological understanding of the case. The authors report that the	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		quality of case analysis/assessment and case formulations were judged to be low in 81% of cases in the intensive sample and they comment that the SCRs '… talk of the need for comprehensive assessments that should be more holistic and coherent in order to paint a full picture. There is a demand that practitioners should look for patterns in the evidence and these patterns should be the subject of a systematic analysis' (p67).	

2. **Cleaver H and Walker S (2004) From policy to practice: the implementation of a new framework for social work assessments of children and families. Child & Family Social Work 9: 81–90**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To evaluate the implementation of the Framework for the Assessment of Children in Need and their Families. To explore impact of the Assessment Framework on practice (the focus of this review). **Methodology** Qualitative study - using audits, postal questionnaires and interviews **Country:** UK. **Source of funding:** Not reported.	**Participants** • Children and young people – aged over 10. • Parents and caregivers. • Professionals/practitioners. **Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported, child protection in general.	**Narrative findings** Views of social work managers, practitioners and parents on the Assessment Framework (from audit records, questionnaires and interviews) 1. Some children considered to be at risk of significant harm received no in-depth assessment of their needs and circumstances, suggesting that information gathered during the initial assessment may not always have been used to inform social work decision making. Social work managers indicate that in some cases the decision to initiate an in-depth assessment was dictated by the availability of services or the legal duties placed on the organisation rather than on the developmental needs and circumstances of children. The absence of in-depth assessment may reflect an organization-led approach to decision making. 2. Social work managers held more positive views than practitioners in the increased involvement/part-	**Overall assessment of external validity** ++ **Overall assessment of credibility (internal validity)** - **Overall score** - Limited methodological details on postal questionnaires (what questions asked, response rates) and interviews (how and where conducted, what questions asked, specifics of par-

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** 1. Postal questionnaires sent to 24 English councils involving 216 social work practitioners, 93 social work managers and 153 professionals from other agencies returned completed questionnaires. **Assessment tool** The UK Framework for the Assessment of Children in Need and their Families. It promotes a holistic, multi-agency approach towards the assessment of children in need, considering the full range of children's and family's strengths as well as needs and difficulties, including the wider environment and circumstances in which they live. This conceptual framework is based around three domains: the child's needs, the capacity of parents or carers to respond appropriately to those needs and family and environmental factors. The Assessment Framework identifies three stages in the assessment process: Referrals, Initial assessment, core assessment.	nership with families and children in assessment because managers found that this involvement led to a. a more transparent and accountable relationship with the family; b. a more focused approach to assessment; c. increased consultation with the family; d. discussing issues where parents and professionals disagree. 3. Social work practitioners reported that the assessment records had hampered the involvement of families due to a. the records being perceived as bureaucratic tools to regularize social work practice. b. appearance and language of the records was not family friendly. c. practitioners unfamiliar with the style and content of the records. This suggests the assessment may necessitate a significant change in the practitioners' recording practice, which can be supported by relevant training at all levels of the organisation to understand the purpose of the assessment records, and the structure and content of the tool. 4. Parents reported satisfaction with the assessment process as they felt consulted and involved in all stages of the process from referral to assessment. They also reported a shared perspective on the difficulties they were facing, an involvement in the choice, development of and commitment to the plan and its fruition. 5. Interview data from children not reported. **Specific impact on social work practice** For social work practitioners:	ticipants and interviewer etc.). Incomplete outcome data: Children were interviewed but no data were presented. The authors stated that 'Because the authors were responsible for some of the research that informed the Assessment Framework and were involved in the development of the assessment records, this study does not purport to be an objective evaluation' (p83).

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Outcomes measured: Satisfaction with services, barriers and facilitators.	1. Increased workload as seeing the child as part of their initial assessment had not been standard practice for social worker, resulting in social workers spending more time with children and families during the process of assessment than they had previously. 2. Anxiety about their ability to analyse the information they collected during the assessment, and collaborative working with colleagues from other agencies. For social work managers: 3. Managers reported an improvement in the quality of the assessment in record keeping, which in turn gave them greater confidence in their own decision making and planning for the children. **Impact on interagency practice** For social work practitioners and professionals of partner agencies 1. Collaborative working between agencies over assessments had increased due to a. information recorded now better structured; b. a more holistic understanding of the child's needs and circumstances; c. greater clarity and responsibility in roles of agencies; d. a greater willingness to share information. 2. Facilitators: a. the involvement of families that the Assessment Framework b. Relevant joint training between professionals from social services and staff from other relevant agencies, adopting a flexible approach that responded to the needs of practitioners and managers. 3. Barriers that hampered collaborative work: a. a lack of agreement over the definition of children in need; b. communication failure between agencies; c. lack of resources identified as necessary by the assessment; d. increased paperwork in assessment in a structured and systematic way without an adequate electronic	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		recording system; e. general difficulties related to introducing a new system.	

3. **Devaney J, Bunting L, Hayes D et al. (2013) Translating Learning into Action: An overview of learning arising from Case Management Reviews in Northern Ireland 2003-2008. Belfast: Queen's University Belfast**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The aim of the report is to 'present key learning from the first 24 case management reviews commissioned and completed [in Northern Ireland] between the commencement of the current process for case management reviews in 2003, up until the end of 2008' (p17). **Methodology:** Other. Analysis of Serious Case Reviews - analogous to thematic analysis of multiple case studies - therefore use qualitative study critical appraisal tool. **Country:** UK. **Source of funding:** Government.	**Participants** • Children and young people. Case Management Reviews concerning children and young people who have died or been seriously injured, and abuse or neglect is known or suspected to have been a contributing factor. • Caregivers and families. • Professionals/practitioners. **Sample characteristics** • Age - Age of index children who were subject to Case Management Review at time of index event were as follows: Under 1 year - 29% Between 1 year and 5 years - 17% Between 6 years and 10 years - 4% Between 11 years and 15 years - 33% 16 years and above - 17% n=24. • Sex - Gender of index child: Female - 54% Male - 46% (n=24). • Ethnicity - All index children were White and had been born in Northern Ireland. • Religion/belief - Not reported.	**Narrative findings** Aspects that hinder effective assessment 1. Reactive responses to critical incidents rather than proactive continual identifying of needs, i.e. crisis situations are treated in isolation rather than looked at in a timeline of events in the child's life. 2. Holistic assessments of risk were not routine. 3. In some cases, no evidence of a systematic comprehensive assessment of need being conducted at any stage in the agencies' involvement with the family. 4. In some cases, a lack of multi-agency working particularly sharing information and completing assessments together. Consequently, no 1 agency has a full picture of the child. 5. Lack of in-depth assessment and analysis of the information gathered by professionals. 6. In one case, 'the assessment on file amounted to little more than a list of family members' (p54). 7. 'No evidence of an effort being made to engage with other key members of the family such as the child's father and background information on the mother's own family and social history were absent' (p54). 8. Limited availability of information relating to the child's parents and significant others.	**Overall assessment of external validity** + **Overall assessment of credibility (internal validity)** + **Overall score** + Overall, there is a lack of description of how thematic analysis was undertaken.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Disability - Of sample of 24 children: - 8 had a mental health disability - 8 had intellectual disability or ADHD - 6 participated in drug or solvent misuse - 5 participated in alcohol misuse - 3 had a physical illness - 1 had a physical disability - 1 had a sensory impairment. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Types of abuse reported under 'indicators of concern' - Family history of child neglect identified in 7 cases - Family history of child emotional abuse identified in 2 cases - Family history of child physical abuse identified in 6 cases - Family history of child sexual abuse identified in 11 cases. • Looked after or adopted status - Looked after status: Looked after at time of index event - 21% Previously looked after - 17% Never looked after - 62% n=24. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** 24 Case Management Reviews analysed.	9. Assessments that are done bureaucratically rather than a genuine assessment of need/risk: 1 example provided was a young person's case was closed and the assessment of need took place after this closure indicating 'this was merely a paper exercise rather than a concerted effort to address identified concerns' (p54). 10. In 1 case, no comprehensive history of the needs of the mother or wider family were conducted despite a long parental history of mental health, alcohol abuse and psychiatric detention. 11. Non-engagement not always highlighting potential risk, in some cases they were closed consequently. 12. Little analysis by professionals to get to the root cause of child's needs and risk. Aspects that support effective assessment 1. Quality of assessment, from these findings it is not clear what quality looks like other than assessments being effective, timely, comprehensive and child focused. 2. Multi-agency working especially communicating child's needs between different professionals. 3. Early identification of needs.	

4. **Horwath J (2005) Identifying and assessing cases of child neglect: learning from the Irish experience. Child and Family Social Work 10: 99–110**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To establish how social workers assess cases of child neglect and to explore with the practitioners and their managers both their perceptions of their practice and factors that impact on practice. **Methodology:** Qualitative study. 1. Case file review and analysis 2. Questionnaires and focus groups for social work practitioners. **Country:** Not UK. Ireland. **Source of funding:** Not reported.	**Participants** • Professionals/practitioners. Social workers in child care (practitioners and managers). **Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Child neglect. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** 1. Case files analysis: n=57 2. Questionnaire surveys (n=40) and focus groups (n= 5 involving 34 social workers).	**Narrative findings** Case files audit (n=57): Characteristics - 130 children aged 0–18 yrs; 62 referrals (40% anonymous referrals from other professionals, 18% from family members); main concerns of referrers were acts of omissions by carers, alcohol use and home conditions). Findings of case file audits (n=57) explored with data from questionnaires (n=40) and 5 focus groups (n=34 staff social workers [SW]). A. Responding to referrals 1. Contacting other professionals Issues with contacting busy professionals, speedier contact via phone calls and clear explanation of what information needed from the professionals a. Although 92% of SWs said that their decision making was influenced by information obtained from other professionals. However, It was unclear as to why certain professionals were contacted while others were ignored b. SWs experienced frustration trying to make contact with busy professionals to ask if they have 'concerns' about the child, leading to contacting only those professionals who could be contacted easily, resulting in decisions being made without the relevant information from all the professionals who knew the family. c. Contact with other professionals were made by sending letters rather than telephoning [note from reviewing team – this may reflect age of study and may no longer be relevant]. Also the nature of 'concerns' was not well understood by the professionals contacted. 2. Contacting the child and family Issues with assessing child neglect issues without communicating well with and seeing children a. Although 49% of SWs	**Overall assessment of external validity** ++ **Overall assessment of credibility (internal validity)** + **Overall score** + Limited methodological details of questionnaire surveys and focus groups. Study was conducted over 10 years ago and practice would have changed since.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		said that their decision making was influenced by communication with the child. Review of case files showed the lack of meaningful communication with children about their lives. Few spoke to children to ascertain their views and feelings, despite when children made very explicit comments about their needs but these were not taken seriously by the social workers. Examples, when child expressed fear of her father, this was ignored and she was interviewed in his presence. b. Some children were not physically seen by SWs, reasons being the children were out or asleep when social worker called. c. Managers recognised need for communication skills and SWs establishing a relationship with the child. d. SWs felt workload pressure as preventing them from spending time working with the children. Issues with carers and the nature of the contact: keeping professionals at bay by carers f. Avoiding/cancelling meaningful visits or appointments by carers, their superficial compliance in response to enquiries from a social worker resulted in case closure without a preliminary or in-depth assessment Issues with which family member to contact g. When both parents were involved, only one of them seen, especially when the parent was described as 'aggressive and intimidating'. h. In many cases, the focus of assessment was on mothers, and fathers was not seen. This placed mothers under pressure, making her responsible for care and protection of the children. 3. Making use of material in case files. Issues with using the wealth of information already available in case files a. The ongoing nature of child neglect issues, such as long history of social work involvement, records of previous referrals, assessments and interventions completed ('video over time' versus 'snapshot')	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		on similar problems were not well used to inform current referrals. B. The assessment process Issues with the focus of assessment 1. The purpose of the assessment a. SWs too concerned with incidence of child neglect, rather than on impact of neglect on the child. Assessment of the child's needs was through talking to carers and professionals and not the child. b. SWs established 3 different perspectives towards assessment process: that child neglect has occurred; that any harm/suffering of child be assessed; and that the impact of harm on the child be assessed. The case files highlighted that in many cases there was no evidence of practitioners assessing the potential harm and impact of the harm on the wellbeing of the child. c. Workloads, systems and resources influenced the team approach to child neglect and can result in short-term intervention. Some focused on the investigation of risk/harm to the child; some focused on keeping the child safe and meeting their needs; and some focused on assessing the impact of neglect on the child, the parenting capacity and parenting environment. d. Approach also influenced by team culture and team leader. 2. Assessing parenting capacity and parenting issues issues a. In some case files, there was no assessment of parents' capacity to meet the needs of their children. When this was considered, the focus was too much on specific weaknesses around parenting capacity, rather than exploring both parental strengths and weaknesses. A strengths-based approach can build on a family's existing competencies and resources to respond to crises and stress, and enhance and strengthen family functioning. b. The impact on of	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		parenting capacity such as alcohol use and depression and was not explored as to how they affect the parents' ability to meet the child's needs. 3. Extended family and environment Issues a. SWs considered poor housing, furnishing and levels of cleanliness in terms of the health and safety needs of the children. A minority focused on what was socially un/acceptable. Some struggled to make allowances for parents doing their best under difficult circumstances while managers were clear that if the child is suffering significant harm action must be taken irrespective of how sorry workers are feeling for the family. 4. Management and the impact on assessment Issues with influence from team managers, workload and resources a. Most SWs believe that decisions should be made taking account of the views of the supervisors. However, supervisors are under pressure, supervision regularly cancelled or interrupted. When it takes place the focus is on cases in crisis and concerns of child neglect are likely to be marginalized in supervision in favour of more crisis driven cases. 5. Assessment outcomes Issues with unclear decision when files closed a. Some case files closed without provision of services as they could not be completed due to uncooperative carers or no response from professionals. When the decision was made to monitor the case, no information in the files indicated what aspects exactly of developmental needs or parenting capacity should be monitored.	

5. **London Safeguarding Children Board (2011) Final monitoring report: local authority pilots of the London safeguarding trafficked children guidance and toolkit. London: London Safeguarding Children Board**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To 'report the experiences of 12 pilot local authorities and three corresponding police forces in implementing the London SCB trafficked children toolkit form January 2009 to May 2010' (p14). **Methodology:** Mixed methods. **Country:** UK. **Source of funding:** Not reported.	**Participants** • Professionals/practitioners. Described as 12 pilot local authorities and 3 corresponding police forces. Roles and exact numbers of participants/respondents in the research not given. **Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** Twelve pilot local authorities (number of people participating in each authority unclear). Three police forces (number of people participating in each police force unclear).	**Narrative findings** Numbering from report: 4.2 Overall rating of toolkit guidance and tools. Minimal reporting of numerical data. Findings reported as follows: 4.2.2 Views on guidance section of the toolkit: - Almost all local authorities found the first four sections of the guidance at least 'useful' - 'Introduction, Definitions, Principles, The problem of child trafficking - Five local authorities rated these section as a 'very useful' or a 'must-have': The problem of child trafficking, Children at risk of or experiencing significant harm, Particularly vulnerable groups of children - Three local authorities also rated 'Safeguarding and promoting the welfare of trafficked children as a 'must-have'. – Section 7 of the toolkit either not used or rated as 'slightly useful' by more than 1 local authority: Introduction, Actions of professionals and agencies, Local expertise in relation to trafficked children, Safeguarding and promoting the safety of trafficked children, Information sharing, Role of Local Safeguarding Children Boards, Role of specific agencies and services. 4.2.3 Views on 16 tools and guidance items - 'On the whole favourably regarded' - Risk assessment matrix was rated as most useful tool - rated as a 'must-have' by 8 authorities and as 'very useful' or 'useful' by 6 authorities. - Assessment framework for trafficked children triangular diagram was rated as second most useful tool - rated as a 'must-have' by 6 authorities and 'useful' or 'slightly useful' by 6 others - Trafficking assessment form: Rated must-have by 2 authorities,	**Overall assessment of internal validity** - **Overall assessment of external validity** + **Overall validity score** - Very poor information regarding methodology, including how pilot sites were recruited, which individuals participated within those sites, and how results were analysed. Incomplete reporting of results and lack of clarity regarding elements of data collection have contributed to which finding.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Outcomes measured Perceived usefulness of London Safeguarding Children Board trafficked children in identifying and assessing children.	very useful by 7. - Quick guide to levels of intervention and Quick referral flowchart: Least useful, rated by 5 authorities as 'slightly useful or not used'. - Policy and legislation, glossary and acronyms, assessment framework for trafficked children: not used by at least 1 local authority. 4.3 Guidance. It is unclear on what data the following feedback is based. - Guidance perceived by local authorities as comprehensive, information and good to 'dip in to' - Considered bulky, dense and too long - Not all local authorities using guidance - some social workers and teams find it more useful than others - Social workers skip main guidance to use tools in Appendix. - More likely to use a condensed version of the guidance. - Toolkit perceived by local authorities as providing good guidance on identification of young people soon after they arrive. Participants thought that more clarity was needed about what to do when a young person has been in care for some time, or there are ongoing suspicions of trafficking or exploitation. - More guidance required on how to help young people recover from trafficking. 4.4 Risk assessment matrix for children who may have been trafficked It is unclear on what data the following feedback is based. Local authority views were as follows: - Risk indicator matrix a useful tool. - Perception that tool may be particularly useful for those dealing with trafficking less often. - Concern that it may be less useful for identifying victims under age 10. - View that risk indicators need to be kept up to date. 4.5 Trafficking assessment form - Participants reported that assessment form had contributed to identification of trafficked children in half the cases (56 cases were identified during project) - Assessment	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		form seen as 'unnecessary, lengthy, repetitive and additional to the rigid Integrated Children's System (ICS) assessment process' (p21) - Lack of time to complete - Concern that cannot train staff to complete such a 'complex assessment' - View that assessment may not be needed where initial/core assessments already being conducted - Not all pilot authorities knew how to integrate the assessment in to their local ICS systems, meaning that trafficking becomes separate from other safeguarding assessments - Reports that authorities have been using risk assessment matrix rather than full assessment framework, possibly because assessment framework perceived to be long, repetitive and too complex - Local authorities thought that it would be helpful to have a lead child trafficking safeguarding coordinator in each local authority - National Referral Mechanism referrals are being made using only risk assessment matrix. - Reasons for not using full assessment document: pilot authorities have learned from experience which parts they need to use, some sections not considered relevant, children have made immediate disclosures of trafficking, questions seen as repetitive and additional to existing assessments. 4.6 Age assessment - Not relevant to the review question. 4.7 Use of National Register for Unaccompanied Children database (NRUC) in assessment process - Five of the 12 pilot authorities' used NRUC as part of assessment and found it useful. 4.8 Impact of trafficking assessment process on child Again, unclear on what part of the data collection these findings are based. - Participants though that multiple assessments might confuse a child and lead to only a partial disclosure. View that information	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		gathering should be more streamlined. - Participants thought that timescales for assessment do not recognise complexity of disclosure. - Participants thought that professionals should make contact with other agencies at first sign of indicators or suspicions, rather than awaiting full disclosure. - Some pilot authorities reported that assessment processes may 'ring-fence' children and label them as separated, trafficked etc. View that initial assessments needed to be able to be altered as necessary.	

6. Ofsted (2014) In the child's time: professional responses to neglect. Manchester: Ofsted

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To explore the effectiveness of arrangements to safeguard children who experience neglect, with a particular focus on children aged 10 years and under. The report draws on evidence from 124 cases and from the views of parents, carers and professionals from the local authority and partner agencies. **Methodology:** Qualitative study. Interviews and discussions.	**Participants** • Caregivers and families – Parents. • Professionals/practitioners - A wide range of professions including the police; health visitors; housing professionals; teachers and learning mentors; a paediatrician; accident and emergency staff; a GP; a family support worker; children's centre workers; adult mental health staff; and social workers from Cafcass. **Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported.	**Narrative findings** Views of social workers on assessment in cases of neglect: 1. The use of standardised approaches (such as the Graded Care Profile) and comprehensive frameworks supported them to assess risk in neglect cases and to monitor change over time, as these methodologies enabled them to have a clear focus on different aspects of neglect, to apply structure and systematic analysis to very complex situations and to identify key areas of risk. This informed better planning of intervention to support and protect the child. 2. However, not all authorities had adopted these assessment models. If consistently adopted, more likely to improve standards of practice, especially if social workers and managers were trained in using the model and managers were effective in quality assuring the standard of work. The Ofsted Report highlighted: 1. The quality of assessments across authorities was too variable, with half of the assessments did not take	**Overall assessment of external validity** + **Overall assessment of credibility (internal validity)** - **Overall score** - Limited reporting of data collection methods.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Country: UK, 11 local authorities in the UK. **Source of funding:** Government – Ofsted.	• Socioeconomic position - Not reported. • Type of abuse - Child neglect. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** 27 referrers interviewed from a wide range of social care professions; 42 referrals concerning neglect selected at random from a list of referrals received by children's social care over the last 6 months.	sufficient account of the family history and the impact on their current parenting were not always considered. 2. In a small number of cases, effective use of chronologies in making assessments was not routinely completed in all cases. Most focus of assessment was on key events in the life of the family rather than its impact on the child. 3. Chronologies sometimes only put together at the point that the decision is made to initiate proceedings, which is clearly far too late in the process (views of 1 legal advisor). 4. The most effective assessments should consider not only the child's perspective and experiences, but also analysed the long-term prognosis for change and the potential long-term impact on children living with neglect. However, very few assessments addressed all of these factors. 5. Professionals lose their focus as some assessments were too focused on the parents' issues rather than on analysis of the impact of adult behaviours on children. 6. Some assessments were characterised by insufficient consideration of the parent–child relationship, with no consideration of attachment behaviour and a lack of attention to the child's emotional and physical development. 7. There was an evident lack of representation of the child's views, wishes and feelings in the assessment process. 8. Training and support to enable social workers to understand and assess the complex range of children's emotional and behavioural difficulties in assessment reports requires further development, as	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		poor assessments can and do result in children being left at risk of harm or being further harmed.	

7. Platt D (2008) Care or control? The effects of investigations and initial assessments on the social worker-parent relationship. Journal of Social Work Practice 22: 301–15

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To explore initial assessment practice and the effects of coercive interventions on relationships between social workers and parents. **Methodology:** Qualitative study. The use of qualitative interviews were conducted both with social workers and parents involved in each of the 23 case studies. Two parents refused to interview. Data analysis software was used (NVivo) and based on a grounded theory approach. **Country:** UK. Setting is two Urban North cities in the UK.	**Participants** • Caregivers and families - 23 case studies were explored through interviews with families who have been subject to an initial assessment (n=20) or investigation (n=3) with borderline child protection concerns. • Professionals/practitioners - 14 social workers were interviewed who were involved with the above 23 case studies explored in this study. **Sample characteristics** • Age - Pre-birth referrals to 1 16 year old. The spread of ages was fairly even but with more 5–9 years old than over 10s. It is not clear on the age of family members or professionals interviewed. • Sex - Predominantly lone mothers (females) were interviewed, with one lone father. Sex of professional not reported. • Ethnicity - 1 mother was of mixed racial origin and one black. Ethnicity of professionals not reported.	**Narrative findings** Findings conceptualised under three key elements: sensitivity, honesty and straightforwardness, and listening and accurate understanding. Comparisons are made throughout of cases the more coercive (investigation) and less coercive (assessment) framework. 1. Sensitivity 1.1 Aspects that hinder effective assessment Some social workers adopt a policing approach, particularly under S.47, which can impact on the parent negatively and make them feel accused before the assessment is concluded. 1.2 Aspects that help The social worker could demonstrate an understanding of the parental difficulties with sensitivity rather than adopting an accusatory manner i.e. 1 social worker commented that threating posture such as 'you need to sort it out, or else' (p306), is unhelpful. Parents indicated that they appreciate an element of fairness and not jumping to conclusions, rather than appearing patronising or superior. 2. Honesty, straightforwardness and provision of adequate information. 2.1 Aspects that hinder effective assessment Parents feeling that things are being done 'behind their backs'. 2.2 Aspects that help	**Overall assessment of external validity** ++ **Overall assessment of credibility (internal validity)** - **Overall score** - Overall, the study meets most of the quality criteria. However, little consideration of impact of the working context in the 2 case study local authorities. This is potentially highly relevant, given that there may have been idiosyncrasies in assessment practices in the two sites. Relatively small number of investigations (n=3) compared

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Source of funding: Not reported.	• Socioeconomic position - The majority of families were living on low incomes. **Sample size** Total 23 case studies involving parents/families and their 14 social workers. **Assessment tool** Comparisons were made between more coercive (S47 investigations) and less coercive (S17 initial assessments).	In one case, a social worker explained that she was up-front with the family, keeping them informed and reassured. A key aspect is to do what you say you are going to do. One parent commented that if the social worker were open with them, they would be open with the social worker, thus highlighting the importance of reciprocal relationships. 3. Listening and accurate understanding 3.1 Aspects that hinder effective assessment It was alluded to by various participants that if a parent/family had a previous negative experience with social services, this was carried through to the next worker. Listening and understanding were highlighted to be a key element in assessment process, however 'the need for workers to take a different perspective on some things they were told presented a dilemma that prevented them from accepting everything the parent said' (p310). The different assessment process i.e. initial assessment vs. S.47 was seen to put additional strain on the relationship between social worker and parent. For example, one 1 social worker commented on completing an initial assessment, there is more time to establish relationships however with an investigation, it is fast and families are sometimes hearing difficult things. Conversely, the paper concludes that the formal status of the case might not be the issue, but 'whether the social worker is in a position of enquiring into reported concerns, irrespective of the label investigation or initial assessment ... the data supported this point' (p311). In addition, the paper found that it the formality of the worker had a stronger impact on parents and skilled workers can develop good working relationships regardless of the procedural context. 3.2 Aspects that support effective assessment	to initial assessments (n=20) on which to form a basis of comparisons.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		The social worker would need to be clear about their statutory duties and powers in the assessment process. Listening was seen as an important 'tool of the trade', 1 social worker commented on the 'importance to listen, value and understand what the parent is saying, but that does not necessarily mean it has to be believed completely' (p309).	

8. Rigby P (2011) Separated and trafficked children: The challenges for child protection professionals. Child Abuse Review 20: 324–40

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To scope the prevalence of child trafficking, profile children and identify factors that facilitate or hinder intervention. **Methodology:** Qualitative study. **Country:** UK, Scotland (Glasgow). **Source of funding:** Not reported.	**Participants** • Children and young people. Case files of 75 unaccompanied asylum-seeking children. • Professionals/practitioners - 16 frontline professionals working with separated or trafficked for children for >3 years. **Sample characteristics** • Age - Children and young people: Aged between 12 and 17 Professionals: Not reported. • Sex - Children and young people: 38 females and 37 males Professionals: Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported.	• **Narrative findings** Identification and assessment Factors that hinder assessment: - Challenges of working with differing cultural experiences, trauma and fear - Challenge of dealing with ongoing grooming relationships with traffickers - Case file analysis showed that initial identification and child protection assessments were 'largely absent' (p333). - Background and journey information is difficult to corroborate - therefore difficult to identify whether trafficked - There are limited links with international agencies who could help with corroboration - Due to lack of disclosure/clear evidence of exploitation professionals use 'indirect indicators of trafficking' (p333) for assessment. - Little understanding of how indirect indicators should be incorporated in to the assessment process. - The distinction between 'trafficking' and 'smuggling' (those who are forced versus those who are consent) was found to be problematic and confusing for workers.	**Overall assessment of external validity** + Assessment is a subset of overall research question. **Overall assessment of credibility (internal validity)** - **Overall score** - Not clear which data were gathered via interview, and which via focus group. Analysis methods unclear. Relatively little reference to, or presentation of, primary data gathered.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Socioeconomic position - Not reported. • Type of abuse - 16 (21%) of the unaccompanied asylum-seeking children in the case file sample were categorised as having been trafficked. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - All children in case file sample (n=75) were unaccompanied asylum-seeking children. 21% were deemed to have been trafficked. **Sample size** Total 75 unaccompanied asylum-seeking children. Total 16 frontline professionals.		

9. Robertson AS (2014) Child welfare assessment practices in Scotland: an ecological process grounded in relationship-building. Journal of Public Child Welfare 8: 164–89

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To understand relational approaches related to child welfare risk assessment. Study has 3 research questions, 1 of which is relevant to our review question: 1. What are Scottish child welfare experts' views of key relationship	**Participants** • Professionals/practitioners -Professionals working with children at risk of, or experiencing, abuse and neglect. **Sample characteristics** • Age - Not reported. • Sex - 11 women, 2 men. • Ethnicity - Not reported. • Religion/belief - Not reported.	**Narrative findings** Key question: What are Scottish child welfare experts' views of key relationship characteristics for working with parents during the assessment process? 1. Listening, remaining calm, persistence - 1 respondents described responding to parents using a gradual, persistent process, allowing time for angry parents to consider concerns and accept that there may be a problem. Another stressed the importance of listening, and that GIRFEC's assessment approach sup-	**Overall assessment of external validity** ++ **Overall assessment of credibility (internal validity)** - **Overall score** -

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
characteristics during the assessment process? Two questions not relevant: 1. How do new, Scottish child welfare policies and specifically GIRFEC address assessment and child wellbeing? 2. What are Scottish child welfare experts' practice experiences with Getting It Right For Every Child (GIRFEC)? **Methodology:** Qualitative study. Multi-methods case study design using 2 Scottish councils. Multi-method data gathering including document review, purposively sampled interviews with child welfare professionals. **Country:** UK, Scotland. **Source of funding:** Other - Conducted as part of PhD thesis.	• Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** Review of 46 documents. Interviews with 13 participants. **Assessment tool** Getting It Right for Every Child (GIRFECT) is a child welfare framework, 'that emphasizes the responsibilities of local communities for caring for all Scottish children' (p173). Led to the development of new assessment concepts, for example the 'My World Triangle'. No further information provided about GIRFEC or My World Triangle.	ported this. Another respondent emphasised importance of persistence, even where there is a poor relationship with families: 'It really is about persistence and maybe just accepting that you may not have a great relationship but you can do the work anyway. You can look at the people you have around you ... maybe health can do that piece of work ... You can get voluntary services in so that they are actually doing the direct work but the social worker has an overview and case management ... Look and see who gets on with them the best' (Direct practice social worker, p179). 2. Clear communication - Importance of honest, clear and transparent communication, including social worker communicating to parents the nature of concerns. Important in providing a framework for families about what to expect from being involved with child welfare services. 3. Confidentiality, the boundaries of information sharing - Importance of being clear about how information will be shared, and with whom. This can be done via a formal consent letter. 4. Trust and engagement - Engagement involves building trust through listening, honesty, confidentiality, helping the families to set clear goals. Helping the families to see the links between these elements, and giving hope that they can change their circumstances. 'There are really no shortcuts to early engagement of parents ... Engagement is a dynamic process not just a series of administrative steps' (Child welfare administrator, p181).	Unclear how thematic analysis of semi-structured interviews has been conducted to arrive at 4 themes of relationship characteristics that support the assessment process. No description of analysis process, whether any qualitative analysis software was used.

10. Selbie J (2009) Health visitors' child protection work: exploratory study of risk assessment. Community practitioner: the journal of the Community Practitioners' and Health Visitors' Association 82(5), 28–31

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To seek health visitors' (HV) opinions on the efficacy of health assessment and screening tools in child protection work: to identify facilitators and enablers in identification and management of risks to children. **Methodology:** Qualitative study: 2 focus groups and 1 1:1 interview. **Country:** UK. **Source of funding:** Not reported.	**Participants** • Professionals/practitioners - Health visitors (HV). **Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not specific - child protection in general. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** Seven HVs (6 HVs in 2 focus groups [3 HVs/focus group] and 1 HV in a 1:1 interview). **Assessment tool** CAF and health assessment and screening tools to identify risk in children.	**Narrative findings** Three main questions asked in interviews: 1. What factors influence identification of risk? 2. What factors influence risk analysis? 3. What factors influence risk management? HVs' views on A. the CAF (Common Assessment Framework): 1. Ambivalence and lack of confidence in the CAF structures - lack of clarity about thresholds between tiers of children services and workload issues. '... I haven't had positive experience of the CAF because the agencies I expected to address particular needs for a family just weren't available …' (HV1). B. Risk analysis 1. A lack of familiarity and confidence with the topic. C. Health visiting skills 1. Co-ordination approach, with documentation and work with other agencies to deliver services: 'a boy with behavioural problems in nursery ... I sought advice from the GP, who referred to the CAMHS … then referred to Sure Start for family support … Sure Start referred it back to me ... gone round and round with no one accepting responsibilities' (HV2). 2. Establishing a good working relationship with families: '... the family has to be aware that you're there in a supportive capacity to enable you to do a thorough assessment ...' (HV4). 3. Skills in interaction with families: '... how I can best present myself and my service … because a lot of families are very suspicious and resistant to engage with you … so it's "How so you sell the whole package!" really ...' (HV6).	**Overall assessment of external validity** ++ **Overall assessment of credibility (internal validity)** + **Overall score** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		4. Different skills mix: '... to be much more informal, and can lead to a lot of things being said which are unlikely to come out with a health visitor ...' (HV2). 5. A trusting and collaborative relationship with the family enables honesty and promotes assessment work. 6. Observational skills: '... if it's a first time visit to the family and there's no history ... you've got to rely on your observations and what they're telling you and pick up other cues…' (HV5). 7. Effective communication skills with families (time to listen) and other professionals an important factor on risk identification, analysis and management: '...I got a better understanding of the situation and I felt more empathy … making an effort to listen to her (mother), I could relate to her better, and she could relate to me better ...' (HV7). 8. Poor communication impacts negatively on risk assessment: '... if a family is well known by a different agency, and you haven't got that information ... that can hinder ... There may be things that are alerting you to certain risk factors ...' (HV8). D. Commitments to families 1. HV highlighted the importance and need of professional commitment to families for long-term assessment work: ' ... With a cursory visit, you probably wouldn't get the whole picture, or find out what had happened quite recently with regards to violence in the family ...' (HV1).	

11. Sen R, Lister PG, Rigby P et al. (2014) Grading the Graded Care Profile. Child Abuse Review 23: 361–73

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The aim of the study is to explore the introduction of the Graded Care Profile, a	**Participants** • Caregivers and families. Total sample unclear (not clear whether some individuals involved in more than one	**Narrative findings** An initial finding of the study was that practitioners were using the GCP 'considerably less' than authority	**Overall assessment of internal validity** -

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
tool used in the assessment of child neglect, in one Scottish local authority. **Methodology:** Mixed methods. Mixed method with 8 different data sources spanning from June 2008 – December 2010: 1. Initial data sought by local authority from 44 practitioners who had used the GCP via questionnaire and interviews; 2. Two focus groups with a subset of above with 14 practitioners; 3. 56 practitioners interviewed via telephone; 4. Semi-structured interview with parents who had previously had the GCP used with them (n=4); 5. Semi-structured interview with practitioners who used the GCP with the above parents; 6. Observation of how the tool worked with three families; 7. Follow up interviews with parents	form of data collection). Data collection activities with parents was via semi-structured interviews (n=4), observation (n=4) post-observation follow up interviews (n=2). • Professionals/practitioners - Total sample unclear (not clear whether some individuals involved in more than one form of data collection). Numbers of practitioners involved in each form of data collection were as follows: questionnaires (n=22), follow-up interviews to questionnaire (n=8), focus groups (n=7), telephone discussions (n=56), semi-structure interviews with workers of involved parents (n=4), post-observation follow-up interviews (n=2). **Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported.	managers had thought. This was reflected in the research design, which aimed to engage with those who had not used the tool, to explore why, as well as those who had. Findings about the GCP are conceptualised under four key headings: user friendliness, the GCP as an assessment tool, parental engagement with the GCP, and the final score. 1. User friendliness 1.1 Aspects that hinder effective assessment Telephone interviews with practitioners noted that the length of time to complete a GCP was seen as a barrier. • The language in the GCP was identified as a barrier for both professionals and parents. This was compounded by the following issues: Questionnaire: - 23% of respondents reported their own understanding of the GCP was a barrier to using effectively - 36% of practitioners reported that the language in the GCP was a difficulty when using with parents - 41% of practitioners found that parental understanding of the GCP was a challenge. Interviews with practitioners: Some practitioners alluded to the challenges with the GCP as going beyond the wording but to cultural assumptions underpinning the tool with one professional commenting, 'it has a real middle class feel to it … the language in it and some of the views about good parenting' (p366). Interviews with parents: The language of the GCP was not a mentioned but featured more predominantly in the observation.	**Overall assessment of external validity** ++ **Overall validity score** - Very little information given regarding data collection or analysis for any of the data collection methods. It is not always clear what data source findings are based on.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
(n=2); 8. Follow up interviews with practitioners (n=2). **Country:** UK, Scotland. **Source of funding:** Other - BASPCAN grant award.	• Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Assessment tool** The Graded Care Profile (Srivastava and Polnay 1997) is a standardised framework for assessment of neglect. It breaks care down in to 4 domains comprising physical care, safety, love and esteem. Each item has a 5-point scale with descriptors for each point on the scale. The scale is completed based on observation and parental self-report.	Observation with practitioners and parents using the GCP: - In two observations, it was evidenced that parents struggled with certain wording such as 'bilateral but overtures more by carer'. This was seen to be overcome by practitioners rephrasing. - Although, practitioners were unsure of how much they should clarify within a standard tool. - With reference to the time taken to use the GCP, this was observed to be between 30 minutes – 3 hours. During the observation, there were interruptions and finding the time to complete the GCP could be challenging. 1.2 Aspects that help • Eighteen out of 20 (82%) of practitioners who completed the questionnaire found the GCP a useful tool for assessing neglect. However, it should be noted that these were all practitioners who were using the GCP, so might be expected to find it more user-friendly. 2. The GCP as an assessment tool 2.1 Aspects that hinder effective assessment • The tool was seen as 'very very subjective', which raises questions about the neutrality and potentially is more a 'value judgment' asset. • According to a few professionals, the GCP might not be 'hugely accurate … the choices … are pretty specific, so there isn't a huge amount of leeway' and additionally, the graded elements of care are judged on what the parents say, rather than through observation. • In an interview, 1 parent disagreed with the social workers grading on	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		the GCP, commenting that they are not always present 24/7 – 'he [social worker] doesn't see it all does he?' 2.2 Aspects that help • The GCP steers discussion by breaking down parental responsibilities. This was highlighted in a focus group where 1 professional commented on the tool being 'invaluable in giving evidence' in the levels of care with a case she managed. • According to some practitioners, the GCP highlights areas of support, i.e. where professionals could be spending more time with families. 3. Parental engagement with the GCP 3.1 Aspects that hinder effective assessment • In one case 'the completion of the GCP was based entirely around what the parent said they did, underpinned by gentle probing' (p369) despite the social worker not having visited or observed the parent first hand as they had recently moved. • The GCP can cause contradictory opinions between social worker and parents if they disagree upon the grade, this was seen to impact on the use of the GCP and in one case a parent abruptly ended the assessment process. 3.2 Aspects that help • Two out of 7 parents agreed and were positive about their overall experience, 1 mother commented that she was glad to have had the GCP. 4. The final score 4.1 Aspects that hinder effective assessment • The GCP was seen to not be the root of disagreements between parent and social workers but it did crystallise underlying issues. • The final score was not always predetermined by the GCP, but compounded with other factors, i.e. in 1 case the GCP prompted a	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		dialogue between mother and social worker rather than diagnosing. 4.2 Aspects that help – none reported.	

12. Vincent S and Petch A (2012) Audit and Analysis of Significant Case Reviews. Edinburgh: The Scottish Government

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The aim of the study was to: '… provide key baseline data on the profile, numbers and emerging themes from Significant Case Reviews conducted in Scotland since 2007, and make conclusions and recommendations about the nature and characteristics of factors which can lead to a Significant Case Review, lessons that can be learned both locally and nationally and implications for both policy and practice' (p30). **Methodology:** Qualitative study. **Country:** United Kingdom. Scotland.	**Participants** • Children and young people. The report is based on an analysis of 56 Significant Case Reviews and 43 Initial Case Reviews conducted after 2007. A Significant Case Review is conducted when a child dies and abuse or neglect is identified as a potential factor; if the child or their sibling was on the Child Protection Register (regardless of whether abuse or neglect is suspected as a factor in the death); if the death was accidental or by suicide; if the child was allegedly murdered or died because of a violent act or reckless conduct; or if the child was looked after. Significant Case Reviews are also carried out in cases of significant harm or risk of significant harm as a result of one of the categories of abuse and neglect specified in 'Protecting Children – A Shared Responsibility: Guidance for Inter-Agency Co-operation'. In addition, there must be serious concerns regarding professional and service involvement in the case. An Initial Case Review is conducted to determine	**Narrative findings** Data extracted from Chapter 3 – 'Practice themes'. The report states that some reviews raise a lack of focus on the child as an issue, noting that in cases involving infants, professionals sometimes focused too much on the needs of the parent at the expense of those of the child. One review stated that: 'Although all practitioners sought to deliver effective services to M and Baby C, the Review found that the child's best interests were ultimately lost sight of in the overall lack of an effective holistic assessment of M's parenting ability' (Details of review unclear, quoted on p64). This is also identified as a factor in cases involving domestic abuse, substance misuse, or mental health, as practitioners working in adult services tended not to focus on the needs of the child. 'The impact of these issues on the health and welfare of the child as a consequence of the actions and lifestyle of the parents was not always fully considered' (p 64). The authors also highlight a tendency amongst some practitioners to take explanations for injury at face value, reporting that 1 Significant Case Review had found that no attempts to verify these had been made. When the same child had not been seen for some time practitioners had failed to consider the possibility of visiting the child in school. The report states a professional tendency to be overly optimistic was another common theme amongst reviews. One review is	**Overall assessment of external validity** + Information on assessment is part of a broader study. **Overall assessment of credibility (internal validity)** + **Overall score** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Source of funding: Government - Scottish Government.	whether a Significant Case Review should be conducted. **Sample characteristics** • Age - Child - Unborn n=2 (3%); under 1 year n=21 (30%); 1–4 years n=18 (26%); 5–10 years n=5 (7%); 11–15 years n=19 (27%); 16 years and over n=5 (7%). Mother - 31 reports did not record the age of the child's mother. Where this information was provided, the details were - 20 – 29 years n=9 reviews; 30 - 39 years n=13 reviews; 40 and over n=3 reviews. Father - 40 reports did not record the age of the child's father. Where this information was provided, the details were - Under 20 years (17 years) n=1 review; 20–29 years n=6; 30–39 years n=5; 40 and over n=4. • Sex - 13 reviews did not record the gender of the children or young people who were the subject of the review and in 2 cases the child had not yet been born. In those reviews which did provide details on gender 59% (n=33) focused on males and 41% (n=23) focused on females. • Ethnicity - Only 2 reviews recorded details of ethnicity and both children were described as White Scottish. In a number of other cases children and caregivers/families were recorded as speaking languages other than English.	quoted as finding that '…so much is recorded as pressures for this family one wonders if the writer is seeking for any evidence of strength to balance these. The danger of this approach is that it is strengths based and potentially underpins a rule of optimism leading to a distorted analysis of impact on or risk to the child ... the welfare of this mother and baby were compromised as a result' (Details of review unclear, quoted on p66). The authors note that some practitioners were unable to 'see or listen' to the child which meant that they missed clear signs of risk and '… did not explore the reasons why the children had run away or consider that the challenging behaviour they were exhibiting might be due to sexual abuse' (p66). Similarly, one review noted that practitioners had failed to consider the role of child's father or siblings in the family, instead focusing on the child and mother. The authors therefore recommend that all family members with an active role in the child's family should be assessed. It is reported that a number of Significant Case Reviews concluded that the initiation of formal child protection procedures had been delayed by practitioner inability to view the case/child holistically (particularly in cases where there was a history of injuries). The authors note that whilst 'no further action' responses may have been procedurally correct and appropriate for the individual agency at that time, if the case had been considered more comprehensively '… concerns may have been escalated to child protection or further assessment may have been undertaken which may have resulted in a different outcome' (p 68).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Religion/belief - The report does not state whether reviews included details on religion/beliefs. • Disability - The report states that none '… of the children in this study were recorded as being disabled but the Significant Case Reviews referred to a number of health problems' (p42) The report does not specifically state whether the reviews included details on the disability status of parents or caregivers, however four cases appear to have involved a parent or parents with a learning disability. • Long term health condition - The report does not state whether reviews included details on long term health conditions. • Sexual orientation - The report does not state whether reviews included details on sexual orientation. • Socioeconomic position - The report does not specifically state whether reviews included details on socioeconomic status. • Type of abuse - Fatal cases - The deaths of children in these cases were attributed to overdose/drug intoxication n=5 reviews; Sudden Infant Deaths/Sudden unexpected deaths in Infancy n=4 reviews; suicide n=3 reviews; natural causes n=3 reviews; infant sleep related deaths n=3; non accidental injury n=2 reviews; child suffocated after the mother fell asleep	The authors found that assessment was discussed specifically in over half of the Significant Case Reviews they studied. Issues identified by reviews included: inadequate exploration of the impact of parental drug misuse; no consideration of the risk arising from domestic abuse; failure to recognise the accumulation of risk factors or to reassess when new concerns were raised; assessments that were a reactive response to an isolated incident rather than an holistic exploration of underlying issues; a failure to analyse historical information and consider what impact this had on ability to parenting; and a failure to involve the police in assessments (even in cases where there was a history of drug use and offending). In relation to record-keeping, the report highlights that some Significant Case Reviews reported that case records were sometimes too descriptive with insufficient analytical insight and did not always flag specific concerns or include the rationale for particular decisions. One review stated that '... records do not reflect their detailed analysis and reasoning. They do not provide a systematic, comprehensive account of what are considered to be the particular risks as well as the specific protective factors' (Details of review unclear, quoted on p77).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	during breastfeeding n=1 review; homicide n=2 reviews; death related to bullying n=1 review; unexplained injury n=1 review; fire death n=1 review. The cause was unclear in one review (pending further investigation) and 2 reviews did not record the cause of death. Non-fatal cases - Physical injury n=11 reviews; ingestion of opiates (i.e. heroin, methadone, etc.) n=6 reviews; neglect n=2 reviews; sexual abuse n=2 reviews; '… concern for unborn child …' n=2 reviews; '… child cruelty and sexual abuse …' n=1 reviews; neglect and sexual abuse n=1 reviews; looked after child convicted of homicide n=1 review; '… safety in care following a complaint by the young person …' n=1 review (p37). • Looked after or adopted status - Nine reviews involved looked after children; and 12 reviews involved children on the Child Protection Register (no further details provided). • Unaccompanied asylum seeking, refugee or trafficked children - The report does not state whether reviews included details on asylum/refugee status or experience or risk of trafficking. **Sample size** Total 56 Significant Case Reviews and 43 Initial Case Reviews.		

Early help

Review question 9: What is the impact of interventions aiming to provide early help to children and young people identified as at risk of child abuse and neglect? (Prevention of occurrence)

Review question 9 – Critical appraisal tables

1. Barlow J, Simkiss D, Stewart-Brown S (2006) Interventions to prevent or ameliorate child physical abuse and neglect: Findings from a systematic review of reviews. Journal of Children's Services 11: 6–28

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. Research question also included reviews on treatment ('indicated') of child abuse and neglect, but data extraction has focused on targeted interventions only. **Adequate description of methodology?** Yes. 'Two reviewers independently assessed the quality of the identified reviews. Data extraction and critical appraisal were conducted on all included reviews' (p8). Data synthesis mainly narrative due to the heterogeneity of the studies involved in these 10 systematic reviews. **Rigorous literature search?** Yes. 'A computerised search was undertaken of key electronic databases: Medline, Psych Info, CINAHL and Social Science Citation Index. Reference lists were also	**Inclusion of relevant individual studies?** Yes. **Study quality assessed and reported?** Yes. Quality assessment (full score 9) of SRs: Of the 10 SRs on targeted interventions, 4 SRs scored 7/9, 3 SR scored 6/9, 3 SRs scored 5/9. Two reviewers independently assessed quality of identified reviews. Critical appraisal conducted. **Do conclusions match findings?** Yes. Conclusions fair and balanced as authors take into consideration the various limitations of summarising 10 systematic reviews involving 509 studies (non RCTs and RCTs).	**Does the study's research question match the review question?** Partly. Ten of 15 reviews related to targeted interventions (meets our PICO), but 5/15 reviews included were related to 'indicated' interventions for treatment of child abuse and neglect. Data extraction has focused on 10 targeted interventions. **Has the study dealt appropriately with any ethical concerns?** Not reported - not applicable. **Were service users involved in the study?** Not reported. **Is there a clear focus on the guideline topic?** Partly. 5/15 reviews included were related to 'indicated' interventions for treatment of child abuse and neglect. **Is the study population the same as at least 1 of the groups**	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall score:** + Limitations: The quality of included reviews is 'fair' due to methodological flaws, surveillance bias, a lack of consensus about the definition of abuse; double counting in which the findings from individual studies were presented more than once. The authors reports that many studies 'did not provide effect sizes for individual outcomes and individual interventions, instead providing "composite" summaries across a range of outcomes and interventions, thereby precluding the possibility of assessing which

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
searched. Key search terms included "Child abuse", "maltreatment", "neglect" and "injuries". These were combined with a range of terms to identify systematic reviews' (p7). Hand searching not reported. All years for which data was available up until December 2005. Year of publication from 1988.		**covered by the guideline?** Yes – families at risk of abuse and neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes - not explicitly stated, home visiting and parenting programmes likely to be delivered in home and primary care setting. **Does the study relate to at least one of the activities covered by the guideline?** Yes - prevention of child abuse and neglect. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes - parenting skills, child abuse and neglect. **Does the study have a UK perspective?** No - different countries, mainly the USA.	type of intervention is best at producing which outcomes' (p23).

2. Carta JJ, Lefever JB; Bigelow K et al. (2013) Randomized trial of a cellular phone-enhanced home visitation parenting intervention. Pediatrics 132 (Suppl. 2): S167–73

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: RCT.	**Was the exposure to the intervention and comparison as intended?** Yes.	**Does the study's research question match the review**	**Overall assessment of internal validity:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Description of theoretical approach? Partly. Theory for supplementation of intervention with cell phones and text messaging given, but not theory for home-based parenting intervention. **How was selection bias minimised?** Randomised. Study reports randomisation, although method not described. **Was the allocation method followed?** Not reported. **Is blinding an issue in this study?** Part blinding. Participants not blind to treatment condition, but research assistants who gathered the data were. **Did participants reflect target group?** Yes. **Were all participants accounted for at study conclusion?** Partly. Study reports a 77% attrition rate - it is not clear whether this differs across conditions.	**Was contamination acceptably low?** Yes. **Did either group receive additional interventions or have services provided in a different manner?** Not reported. **Were outcomes relevant?** Partly. Risk of abuse and neglect only measured through parenting stress index - no other maltreatment measures, including incidence of maltreatment/CPS reports. **Were outcome measures reliable?** Partly. Study mainly used validated measurement scales, except for PAT (Planned Activities Training) checklist, which is used to assess use of PAT strategies. It is unclear whether this has been shown to be a reliable or valid measure of improvements in parenting. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes.	**question?** Yes. Study is examining the effectiveness of an intervention for families at risk of maltreatment. **Has the study dealt appropriately with any ethical concerns?** Yes. Procedures approved by Institutional review board, and informed consent gained from participants. **Were service users involved in the study?** No. Service users involved as participants but not in design or analysis of results. **Is there a clear focus on the guideline topic?** Yes. Focus is on early help. **Is the study population the same as at least one of the groups covered by the guideline?** Yes. Population is parents at risk of maltreatment. **Is the study setting the same as at least one of the settings covered by the guideline?** Yes. Participants' homes. **Does the study relate to at least one of the activities covered by the guideline?** Yes. Relates to	**Overall assessment of external validity:** + Not a UK study. **Overall validity score:** + It is a significant limitation; given that part of the rationale for the intervention is to maintain involvement; that attrition rates across conditions do not appear to have been monitored. There is also no comparison of characteristics of participants in different conditions at baseline and a relatively short follow-up.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Were there similar follow-up times in exposure and comparison groups? Yes. **Was follow-up time meaningful?** Partly. Follow-up time was 6 months. Other studies (e.g. Guterman et al. 2013) have suggested that this may be too short to observe differences. **Was intention to treat (ITT) analysis conducted?** Not reported. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. No power calculations or expected effect sizes given. **Were the estimates of effect size given or calculable?** Yes: d scores and standardised beta reported. However, d scores only given in text, not in tables. **Were the analytical methods appropriate?** Yes. **Was the precision of intervention effects given or calculable? Were they meaningful?** Partly. Confidence intervals provided for changes in mean scores, but not	early help. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **Does the study have a UK perspective?** No.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	for d or standardised beta estimates.		

3. Dawe S and Harnett P (2007) Reducing potential for child abuse among methadone-maintained parents: Results from a randomized controlled trial. Journal of Substance Abuse Treatment 32: 381–90

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: RCT. **Description of theoretical approach?** Partly. The authors note in the abstract that the Parents Under Pressure programme (PUP) draws from the ecological model of child development but the specific theory is not discussed in great detail. It is noted that families in which one or both of the parents abuse substances often show high rates of child maltreatment but that this is more likely a result of multiple domains of family functioning such as paternal psychology, the family environment and parental psychology rather than parental drug use as the sole risk factor. The authors go on to emphasise the importance of developing interventions which address high risk families, noting that there have been relatively few so far and that findings related to these have been	**Was the exposure to the intervention and comparison as intended?** Partly. The Parents Under Pressure programme is intended to be delivered in weekly sessions over ten to 12 weeks. The authors report that there was a high rate of engagement in the experimental condition although it should be noted that four of the 22 families assigned to this treatment condition, only received seven to eight sessions. Engagement levels in the comparison groups are not reported. **Was contamination acceptably low?** Partly. Contamination rates are not reported specifically but the authors note that adherence to the Parents Under Pressure programme '... was maintained by close supervision of treatment progress ensuring that parent workbooks and treatment plans reflected the formulation of each individual PUP family' (p383). Both	**Does the study's research question match the review question?** Yes. The study aimed to evaluate the impact of the Parents Under Pressure programme on outcomes such as family functioning (including child abuse potential) in families in which a parent was engaged in a methadone maintenance programme. This was compared to standard care and a '... second brief intervention control group ...' (p381). **Has the study dealt appropriately with any ethical concerns?** Yes. Participants volunteered, were informed that the study had three treatment conditions and were told what these were. The study was approved by '... hospital and university human ethics committees' (p383). **Were service users involved in the study?** No. No indication that service users were involved at the	**Overall assessment of internal validity:** + This is a well-designed study but the sample size is very small (n=64) and the follow-up period is quite short. In addition, families in the PUP programme may have received some other services. **Overall assessment of external validity:** + Only awarded a + as the study was conducted in Australia. The Guideline Committee should also bear in mind that the study was conducted with parents who abuse substances which has implications for external validity. **Overall validity score:** + The study seems to have been well conducted on the whole however the short follow-up period

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
mixed. A 2003 study of the Parents Under Pressure programme showed clinically significant improvements on a range of outcomes (although maltreatment is not mentioned specifically) which provided the impetus for the current study. **How was selection bias minimised?** Randomised. Method unclear – 'Participants were allocated to one of the three treatment conditions on the basis of a previously determined randomized order of treatment once eligibility had been confirmed' (p383). **Was the allocation method followed?** Partly. No information on allocation concealment is provided. **Is blinding an issue in this study?** Part blinding. Participants were informed that the study had 3 treatment conditions and what these were. It would not have been possible to blind providers. Post-treatment and 6 month follow-up assessments were conducted by an independent research assistant (no information on who conducted baseline assessments is given).	the Parents Under Pressure programme and the 'brief intervention' were delivered by the same therapists. **Did either group receive additional interventions or have services provided in a different manner?** Partly. Participants in the Parents Under Pressure programme could also receive 'additional case management' outside of treatment sessions. This appears to have included school visits (n=10), accompanied legal visit (n=7), social services liaison (n=4, accompanied child health service visits (n=9 families), accompanied supermarket visits (n=6). **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Partly. All measures had pre-established reliability and validity however with the exception of 1 measure of parental substance abuse they all relied on self-reported data. The authors note this as a limitation of the study and their use of the Child Abuse Potential Rigidity scale may have countered this to some extent in	design stage. **Is there a clear focus on the guideline topic?** Yes. Families in which the primary caregiver was engaged in treatment for heroin addiction. The abstract emphasises the high rates of abuse and neglect which occur in families where one or both parents have substance abuse problems. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Families in which the primary caregiver is accessing substance abuse treatment. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Parents Under Pressure was delivered in the family home; the 'brief intervention' was delivered in the methadone clinic; however it is not clear where standard care was delivered – likely to have been delivered in the clinic. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. The Parents Under Pressure programme tar-	and small sample size are significant limitations; and the study was conducted in Australia.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Did participants reflect target group? Yes. The study targeted parents who were being prescribed methadone who had children between the ages of 2 and 8. This age group were targeted due to evidence showing that parenting interventions are more effective for younger children. Recruitment was through two inner city community methadone clinics. Families were recruited through posters displayed in the clinics. To be eligible, the primary carer needed to be currently being treated with methadone, have at least one child between the ages of 2 and 8 for whom they were the full-time carer and be able to understand and read English. Seventy-seven clinic parents were screened for eligibility who were randomised to one of the three treatment conditions. After randomisation, ten declined to participate (n=3 'brief intervention', n=7 standard care). Prior to assessment, 2 parents were incarcerated and 1 withdrew after the birth of another child. In total. 64 families were assessed at baseline. **Were all participants accounted for at study conclusion?** Yes.	measurements of child abuse potential (they suggest that it is '... less influenced by attempts to present in a more positive light (Milner & Crouch, 1997) with some evidence indicating that those parents who are attempting to present themselves in a more positive light (elevated Faking Good scores) have higher rigidity scores than those whose scores are valid (Carr, Moretti, & Cue, 2005)' (p384). **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. Although incidence of abuse and neglect was not measured. **Were there similar follow-up times in exposure and comparison groups?** Yes. All groups were assessed at baseline, post-treatment or three months, and 6 months. **Was follow-up time meaningful?** Partly. Final assessments were conducted at the 6-month point which is a very short timescale and means that longer term im-	gets family functioning and parenting skills such as non-punitive child behavioural management. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. The study measures child abuse potential, parenting stress, child behaviours and parental substance abuse. **(For views questions) Are the views and experiences reported relevant to the guideline?** Not applicable (not views question). **Does the study have a UK perspective?** No - Australian.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Rates of attrition were acceptable in both the Parents Under Pressure group and the 'brief intervention' group; however the attrition rate was quite high in the standard care group (32%). At final assessment at 6 months post-treatment - Parents Under Pressure: 20 out of 22 participants provided assessment data. 'Brief intervention': 20 out of 23 participants provided data. Standard care group: 13 out of 19 families provided data.	pacts of the interventions are unclear. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. The authors report that there were no significant differences between groups at baseline. They note that although the mean daily dose of methadone did not differ significantly between groups the dose effect did approach significance. **Was intention to treat (ITT) analysis conducted?** Not reported. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. Power calculations and effect sizes are not provided. As a pilot study, the sample size was very small (64 in total) and may not have been sufficient to detect effects. **Were the estimates of effect size given or calculable?** No. **Were the analytical methods appropriate?** Yes. Multilevel linear mixed modelling was used to compare trajectories of the PUP programme and 'brief intervention'		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	group to standard care group. Clinical significance was assessed using changes in scores on the Child Abuse Potential scale from baseline to six month assessment. In addition, a Reliable Change Index (Jacobson and Truax 1991) was calculated for these scores. A score was deemed to be clinically significant if it was greater than 1.96. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes - p values are provided. **Do conclusions match findings?** Yes.		

4. DePanfilis D and Dubowitz H (2005) Family connections: A program for preventing child neglect. Child Maltreatment 10: 108-123

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: RCT with 2 intervention groups; receiving 1 Family Connections for 3 months (FC3) and 1 receiving intervention for 9 months (FC9). **Description of theoretical approach?** Yes. Intervention based on '… principles of prevention science [which] suggest that prevention programs should reduce risk factors and promote protective	**Was the exposure to the intervention and comparison as intended?** Partly. Participants in the 9-month exposure group did not receive a full three times the 'dosage' of those in the 3-month exposure group, due to a 'tapering off' in services over time. Families served for 3 months were provided an average of 1.4 hours per week, and families served for 9 months were provided an average	**Does the study's research question match the review question?** Yes. Study examines the effectiveness of a programme to prevent child neglect. **Has the study dealt appropriately with any ethical concerns?** Yes. Approval by Institutional Review Board and informed consent was sought.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Not ++ as not a UK perspective. **Overall validity score:** - Main methodological limitations are as follows: Group allocation

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
factors (Mrazek and Haggerty, 1994; Schinke et al. 1986)' (p109). **How was selection bias minimised?** Randomised. Randomisation based on pre-generated random assignment table. **Was the allocation method followed?** Yes. **Is blinding an issue in this study?** Blinding not possible. Not possible to blind participants to study condition, as based on duration of time in the programme. However, also does not report whether assessors were blind to treatment condition. **Did participants reflect target group?** Yes. Although 62 families who were eligible were not provided with services, as they did not agree to sign up to weekly meetings. This suggests that data could be skewed in favour of more motivated families. These 62 families do not appear to be included in the intention to treat analysis. **Were all participants accounted for at study conclusion?** Partly. The original sample comprised	of 0.9 hours per week. **Was contamination acceptably low?** Partly. In the discussion, the authors hypothesise that there may have been less difference between the two 'dosages' of intervention than hypothesised, because families in FC3 were referred to community services following case closure. **Did either group receive additional interventions or have services provided in a different manner?** Yes. Parents in FC3 were referred to community services on case closure so may, in effect, have continued receiving services. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Yes. Standardised measures used for all outcomes. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up**	**Were service users involved in the study?** No. Service users were subjects of the study, but not involved in design, interpretation of results and so on. **Is there a clear focus on the guideline topic?** Yes. Focus is on early help. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population is parents and caregivers of children at risk of neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Intervention delivered in participants' homes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to early help. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes - study looks at impact on neglect. **Does the study have a UK perspective?** No – USA.	based on 'dosage' means that it is difficult to interpret main effects of time on the whole sample - would have been a stronger design to have a 'usual care' control group. - Possible selection bias in favour of more motivated families. That is, families who were eligible but did not agree to weekly contacts for up to 9 months were not provided with services and, it appears, not included in analyses. - No analysis of any systematic differences in risk factors between intervention conditions. - Unclear whether 3 months is a valid duration for services. Other authors have commented that even 6 months is relatively brief (Guterman et al. 2013).

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
154 participants, but only 125 participants completed data collection at all three time points. Unclear why this is the case. Missing data points were not imputed, so these 29 participants were effectively excluded.	**times in exposure and comparison groups?** Partly. Both 3- and 9-month intervention groups received a further follow up at 6 months post intervention. **Was follow-up time meaningful?** Partly. In the context of other studies in this review (e.g. Guterman et al., 3 months would appear to be a relatively short duration of intervention). **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Partly. Study reports that '… there were no significant differences between participants assigned to receive FC3 versus FC9 in caregiver age, educational level, income or total number of children …' (p110), but no significance testing is shown to support this. Also, there does not appear to be any analysis of any significant differences in risk factors at baseline. **Was intention to treat (ITT) analysis conducted?** Partly. Appears that ITT analysis was applied to families who were enrolled in the studies, but excluded those who were eligible but not enrolled.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Was the study sufficiently powered to detect an intervention effect (if one exists)? Not reported. No power calculation given. Also no information on expected effect sizes, so not possible to calculate power. **Were the estimates of effect size given or calculable?** No. F values and p values given only - effect sizes not given, but could be calculated. **Were the analytical methods appropriate?** Yes. Analysis of variance with 1 within-groups factor (time) with 3 levels (baseline, case closure and 6-month follow up) and one between-groups factor (allocation to 3 or 9 month condition). Analyses tested main effects of both factors, and the interaction (group by time). **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes – p values given.		

5. DePanfilis D, Dubowitz H, Kunz J (2008) Assessing the cost-effectiveness of Family Connections. Child Abuse and Neglect 32: 335–51

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: RCT with two intervention groups: 1 receiving Family Connections for 3 months (FC3) and 1 receiving intervention for 9 months (FC9). Note: This study is based on same participants as DePanfilis et al. (2005). **Description of theoretical approach?** Yes. Logic model for intervention, linking programme inputs, outputs, short term intermediate outcomes and programme outcomes is provided in figure 1 (p341). **How was selection bias minimised?** Randomised. **Was the allocation method followed?** Yes. **Is blinding an issue in this study?** Blinding not possible. Not possible to blind participants to study condition, as based on duration of time in the programme. Not reported whether assessors were blind to treatment conditions. **Did participants reflect target group?** Yes.	**Was the exposure to the intervention and comparison as intended?** Partly. Participants in the 9-month exposure group did not receive a full 3 times the 'dosage' of those in the 3-month exposure group, due to a 'tapering off' in services over time. Families served for 3 months were provided an average of 1.4 hours per week, and families served for 9 months were provided an average of 0.9 hours per week. **Was contamination acceptably low?** Partly. In original paper's (DePanfilis et al. 2005), in the discussion section, the authors hypothesise that there may have been less difference between the two 'dosages' of intervention than hypothesised, because families in FC3 were referred to community services following case closure. **Did either group receive additional interventions or have services provided in a different manner?** Yes. Parents in FC3 were referred to community services on case closure so may, in effect, have continued receiving services.	**Does the study's research question match the review question?** Yes. Study is examining the effectiveness and cost effectiveness of a programme to prevent child neglect. **Has the study dealt appropriately with any ethical concerns?** Yes. Approval by Institutional Review Board and informed consent. **Were service users involved in the study?** No. Service users were subjects of the study, but not involved in design, interpretation of results and so on. **Is there a clear focus on the guideline topic?** Yes. Focus is on early help. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population is parents and caregivers of children at risk of neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Intervention delivered in participant homes.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Not ++ as not a UK perspective. **Overall validity score:** -

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Were all participants accounted for at study conclusion? Yes. Although in original study - De-Panfilis et al. (2005) - 62 families who were eligible were not provided with services, as did not agree to sign up to weekly meetings. This suggests that data could be skewed in favour of more motivated families. These 62 families do not appear to be included in the intention to treat analysis.	**Were outcomes relevant?** Yes. **Were outcome measures reliable?** Yes. Standardised measures used for all outcomes. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Partly. Both 3- and 9-month intervention groups received a further follow up at 6 months post intervention. **Was follow-up time meaningful?** Partly. In the context of other studies in this review (e.g. Guterman et al. 2013) 3 months would appear to be a relatively short duration of intervention. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Partly. Study reports that 'there were no significant differences between participants assigned to receive FC3 versus FC9 in caregiver age, educational level, income or	**Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to early help. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Study looks at impact on neglect. **Does the study have a UK perspective?** No.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	total number of children' (p340), but no significance testing is shown to support this. Also, there does not appear to be analysis of any significant differences in risk factors at baseline. **Was intention to treat (ITT) analysis conducted?** Partly. Appears that ITT analysis was applied to families who were enrolled in the studies, but excluded those who were eligible but not enrolled. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. No power calculation given. Also no information on expected effect sizes, so not possible to calculate power. **Were the estimates of effect size given or calculable?** No. Effect sizes are reported in a previous study (DePanfilis et al. 2005). **Were the analytical methods appropriate?** Partly. As ANOVA results were reported in previous study (DePanfilis et al. 2005), they are not replicated here. However, this study reports changes in raw scores, along with p values. It is not clear with what statistical test		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	these p values are associated. Similarly, cost effectiveness was calculated by comparing costs to raw score changes. **Was the precision of intervention effects given or calculable? Were they meaningful?** No. No confidence intervals given with regard to costs of changes in scores. Also changes in raw scores reported without standardisation in relation to standard deviation. **Do conclusions match findings?** Yes.		

6. Dishion T, Mun Chung J, Drake Emily C et al. (2015) A transactional approach to preventing early childhood neglect: The Family Check-Up as a public health strategy. Development and psychopathology 27: 1647–60

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: To investigate whether a home-based visitation intervention, the Family Check-Up (FCU) reduces the risk of child of maltreatment by improving the parent-child relationship in low-income US families. **Description of theoretical approach?** Yes. Based on theory	**Was the exposure to the intervention and comparison as intended?** Yes. **Was contamination acceptably low?** Not reported. **Did either group receive additional interventions or have services provided in a different**	**Does the study's research question match the review question?** Yes. Incidence of neglect and parenting quality. **Has the study dealt appropriately with any ethical concerns?** Yes.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Lack of UK focus **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
that more extreme forms of maltreatment emerge from daily conditions and interactions of a neglectful caregiving environment; that it is important to provide treatment for families that show early signs of maltreatment that may 'prevent escalation of problematic parenting' (p1648), that it is important to understand the mediating and moderating mechanisms that link prevention strategies to reduced risk of child maltreatment. It argues that it is likely an overall level of family adversity can affect levels of parenting stress and potential for positive parent-child engagement and that home-based services break down the barrier for stressed families unable to reach out to support services due to lack of transport. **How was selection bias minimised?** Randomised. **Was the allocation method followed?** Yes. **Is blinding an issue in this study?** Part blinding. Assessors blind to group status of participants, mothers aware of allocation. **Did participants reflect target group?** Yes.	**manner?** No. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Yes. Use of 2 validated measures HOME, family adversity index, and items from the COIMP. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Yes. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. **Was intention to treat (ITT) analysis conducted?** Yes. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. Good sized sample of 731. **Were the estimates of effect size given or calculable?** Partly. Significance tests and confidence	**Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Primary caregivers at high risk of neglect **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Own homes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Early help. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Incidence of neglect and parenting quality. **Does the study have a UK perspective?** No. American.	Rationale for data analysis not always clear.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Were all participants accounted for at study conclusion? Yes. 90% retention rate.	intervals provided for family adversity scores, for indirect effects shown in Table 3. **Were the analytical methods appropriate?** Partly. Results analysed using path analysis. Unclear why some data have been used and not others, e.g., measurements of dyadic positive engagement were taken at ages 2, 3, 4 and 5 but only the data for age three years were used in the model. **Was the precision of intervention effects given or calculable? Were they meaningful?** Partly. For final model of indirect effects of intervention. **Do conclusions match findings?** Partly. Discusses further the links between the three outcomes but does not deal with why there was no direct effect seen by the intervention on neglect.		

7. DuMont K, Kirkland K, Mitchell-Herzfeld S et al. (2011) Randomized Trial of Healthy Families New York (HFNY): Does Home Visiting Prevent Child Maltreatment? New York: New York State Office of Children and Family Services

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: RCT. **Description of theoretical approach?** Yes. Approach is	**Was the exposure to the intervention and comparison as intended?** Partly. 9.8% of participants in the intervention condition	**Does the study's research question match the review question?** Yes. Study looks at	**Overall assessment of internal validity:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
based on importance of early experiences in terms of: association between harsh, abusive or neglectful parenting practices and risk of engaging in violence, substance misuse and juvenile delinquency (e.g. Eron et al. 1991) - modelling inappropriate or violent behaviours (Farrington, 1991) - causing biological, neurological or cognitive problems (Feldman and Downey 2004). **How was selection bias minimised?** Randomised. Computer randomisation. **Was the allocation method followed?** Yes. **Is blinding an issue in this study?** Part blinding. Assessors were blind to parent allocation condition. Nature of intervention means that allocation could not be concealed from participants themselves. **Did participants reflect target group?** Yes.	did not receive intervention, but were included in analysis (ITT). **Was contamination acceptably low?** Partly. 2.5% of participants in the control condition erroneously received services. **Did either group receive additional interventions or have services provided in a different manner?** No. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Yes. Outcome measures for child maltreatment comprised administrative data and standardised scales such as the revised parent-child Conflict Tactics Scale (Straus et al. 1998). **Were all outcome measurements complete?** Partly. Parenting measures were collected at Year 7 only - no baseline data. **Were all important outcomes assessed?** Yes.	the impact of an early help intervention on child maltreatment. **Has the study dealt appropriately with any ethical concerns?** Yes. Study was reviewed by the Institutional Review Board of the University of Albany. Informed consent from all participants was sought. However, there is not much consideration of the impact of being in the control group (and therefore receiving no services). **Were service users involved in the study?** No. Service users involved as participants, but do not appear to have been involved in terms of informing the design or analysis of the study. **Is there a clear focus on the guideline topic?** Yes. Study is relevant to early help aspect of guideline topic. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population is caregivers of children at risk of abuse and neglect.	**Overall assessment of external validity:** + Not ++ as does not have a UK focus. **Overall validity score:** + A well-designed study with substantial follow-up time, and good retention rate given this is a 7-year follow up. Key concerns are: the validity of the 2 subgroup analyses (small sample sizes), and why data from Waves 2 and 3 of the study were not included in the analysis.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Were all participants accounted for at study conclusion? Partly. Study has good retention rates (90% at Year 1 follow up, 85% and Year 2 and 80% of baseline at Year 7). However, no analysis of drop-outs - for example whether these were higher in control compared to intervention group.	**Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Yes. Follow-up at 1, 2 and 7 years. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. Intervention and control groups did not differ significantly on any demographic variables at baseline, except that the control group were significantly more likely to have a female target child. This has implications for later measures of externalising behaviour, which tend to be higher in boys. Intervention and control also did not differ on risk scores at baseline, except for the Kempe Family Stress checklist score for overall count of risk items, which was significantly higher in intervention group. **Was intention to treat (ITT) analysis conducted?** Yes - for analysis of administrative data. For assessed measures, not possible to include data for families	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Setting is families' homes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to early help. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Study considers maltreatment outcomes. **Does the study have a UK perspective?** No.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	who did not take part in assessment. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. Power calculations for the study as a whole are not presented. Consideration is given to the power for the 2 subgroup analyses (Recurrence Reduction Opportunity subgroup and High Prevention Opportunity subgroup). For RRO group, effect sizes of 0.25 were detectable (at 0.05 confidence level); for HPO group effect sizes of 0.2 were detectable (at 0.05 confidence level). **Were the estimates of effect size given or calculable?** Yes. Study uses either effect sizes or adjusted odds ratios, calculated via logistic regression. **Were the analytical methods appropriate?** Partly. Analytical techniques largely good, and different analytical techniques used for normally versus non-normally distributed dependent variables. However, unsure about validity of two subgroup analyses (RRO and		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	HPO) which are both based on very small n (<60). **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. **Do conclusions match findings?** Partly. Yes, although reference made to 'sustained change' from Waves 2 and 3 - but data is not presented here.		

8. Green BL, Tarte JM, Harrison PM et al. (2014) Results from a randomized trial of the Healthy Families Oregon accredited statewide program: Early program impacts on parenting. Children and Youth Services Review 44: 288–98

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: RCT. **Description of theoretical approach?** Yes. The study's approach is based on the Healthy Families Oregon logic model (the provision of core services lead to improvements in short term outcomes which result in the long term goals of preventing child maltreatment and increasing school readiness). Core services are identified as parenting education and coaching, identification of indi-	**Was the exposure to the intervention and comparison as intended?** Not reported. It is quite unclear which participants received the intervention. The authors surveyed participants on service utilisation to determine which services families had received home visiting services. This information was then 'augmented' and 'verified' using the records of the programme itself. This led to 're-coding' on responses where it appeared that a family had received the intervention. The authors note	**Does the study's research question match the review question?** Yes. The study aimed to determine the short-term effects of the Healthy Families Oregon program on parenting behaviours, parenting stress and depression, family functioning and child development. The authors note that the scale used to determine parenting stress is associated with higher risk for maltreatment. **Has the study dealt appropriately with any ethical concerns?**	**Overall assessment of internal validity:** - Key study limitations are a lack of baseline data for relevant outcomes, and a lack of clear description of control intervention. **Overall assessment of external validity:** + Only given a + as the study was conducted in the USA. **Overall validity score:** -

vidual family issues which may impede development and parenting, support of healthy child development by home visitors. Short term outcomes are identified as increased parenting skills, reductions in parental risk factors and improvements in child development and health. **How was selection bias minimised?** Randomised. Screening results of eligible families (using the New Baby Questionnaire) were added to a state database (web based) which used a random-number generator to randomise participants. This paper reports on a smaller subset of this sample who were randomly selected by the researchers. No details on this specific process are provided. Analysis of maternal demographic and risk variables showed that differences between the two groups were non-significant with the exception that the intervention group were more likely to report 'family relationship problems'. **Was the allocation method followed?** Not reported. No details indicating allocation concealment are provided. **Is blinding an issue in this study?**	in the limitations section that a '... relatively large percentage of parents who were screened and offered home visiting were never actually served by the program' (p296). They attribute this to feelings amongst families that they did not need the service, or that staff were unable to locate or contact the family after screening for eligibility. In addition, the authors note that there can be local variability in the model and as it appears that families may have been served by one of seven programmes it is difficult to be confident that the interventions received were comparable. **Was contamination acceptably low?** Yes. The authors report that one control participant had received a home visit from the programme but given the method by which exposure data was collected it seems difficult to be confident that this is the only case of contamination. **Did either group receive additional interventions or have services provided in a different manner?** Yes. Families in both groups reported that they had attended parenting classes. **Were outcomes relevant?** Yes.	Partly. There is no detail given on ethical approval but participants gave consent to be screened for eligibility for the programme and to the use of their medical records. The main source of data which the paper draws on is a telephone survey which it is stated a number of families declined to participate in. The study also excluded eligible infants who were 'medically fragile' or had special needs and infants at risk of removal because of a positive toxicology screen at birth or because of immediate safety concerns. Both groups remained eligible for the programme. **Were service users involved in the study?** No. No indication that service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by**	As this study was conducted in the USA and has a number of important methodological problems (e.g. lack of clarity regarding control intervention, lack of baseline data for outcome measures, failure to account for systematic differences in risk factors between treatment and comparison group) means that it is not possible to award a higher overall validity score.

Part blinding. Telephone interviewers were blinded to assignment condition although the authors note that this may have become apparent during interviews. Participants were not blinded and it would not have been possible to blind providers. **Did participants reflect target group?** Partly. The programme enrolled 2664 families in total but no details are provided on the percentage of eligible families who agreed to participate in the programme. The programme targeted first time parents with babies younger than 3 months. Programme eligibility was assessed using the New Baby Questionnaire. This is adapted from the Hawaii Health Risk Indicators Instrument (Duggan 2004). This measure assesses family risk using a range of criteria: mothers under the age of 19, delayed prenatal care, poor engagement with prenatal care services, single parent, depression (measured in mothers using PHQ-2), low education, drug abuse, troubled family relations. Families were deemed eligible if they scored positively for any two risk criteria or if there were substance abuse or depression issues. The majority of partic-	Although it is not clear that the Parenting Stress Index is an adequate measure of maltreatment. **Were outcome measures reliable?** Partly. Most measures collected using valid and reliable scales however they all rely on self-reported data. Data on breastfeeding, developmental screening and child's developmental status were not collected via established scales. **Were all outcome measurements complete?** Partly. Baseline assessments of outcome measures were not conducted. **Were all important outcomes assessed?** Partly. The study did not include any measures which directly assessed child health and development which seems problematic given the short term outcomes detailed in the logic model. **Were there similar follow-up times in exposure and comparison groups?** Yes. Interviews took place around the time of the child's first birthday. **Was follow-up time meaningful?** Yes. The study only aimed to as-	**the guideline?** Yes. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **(For views questions) Are the views and experiences reported relevant to the guideline?** Not applicable (not views question). **Does the study have a UK perspective?** No - USA.	

ipants were white and large percentages were single mothers who experienced financial difficulties. 82% of enrolled families agreed to be contacted for the telephone interview which this study is based on. From the group of families who agreed to be contacted for the researchers randomly selected a sample of n=1604 in order to achieve the target of 800 interviews. The authors then go on to state that families were 'replaced' in the phone survey sample if they '... had no working phone number after multiple attempts (n = 494, 30.8% of those attempted); if they were unable to be reached after 10–20 telephone attempts and two attempts by mail (n = 269, 16.7%); or if they declined to participate in the survey (n= 42, 5.2% of those contacted)' (p291). **Were all participants accounted for at study conclusion?** Yes. Cross-sectional data used.	sess short term impacts so the follow up at child's first birthday seems reasonable. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Partly. Analysis of maternal demographic and risk variables showed that differences between the two groups were non-significant with the exception that the intervention group were more likely to report 'family relationship problems'. This does not appear to have been controlled for. **Was intention to treat (ITT) analysis conducted?** Yes. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. No power calculations or expected effect sizes are provided. The sample size seems sufficient. **Were the estimates of effect size given or calculable?** Yes. Effect sizes using partial eta-squared (continuous outcomes) and odds ratios (dichotomous outcomes) are provided. **Were the analytical methods ap-**		

	propriate? Yes. ANOVA, ANCOVA and logistic regression were used. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes - p values are provided. **Do conclusions match findings?** Partly - On the whole, the authors conclusions do match their more detailed findings, however they begin their discussion by noting that mothers in the intervention group were significantly more likely to read to their infant which doesn't quite seem to tally with their main goals.		

9. Guterman NB, Tabone JK, Bryan GM et al. (2013) Examining the effectiveness of home-based parent aide services to reduce risk for physical child abuse and neglect: Six-month findings from a randomized clinical trial. Child Abuse and Neglect 37: 566–77

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: RCT. **Description of theoretical approach?** Yes. Refers to theoretical basis in ecological and stress theories (e.g. Belsky 1993), and identifies four key mechanisms by which interventions should have effect: 1) child safety 2) parenting skill guidance 3) problem-solving support 4) improving parents' social support. **How was selection bias minimised?** Randomised. **Was the allocation method followed?** Yes. **Is blinding an issue in this study?** Partial blinding. Blinding not possible for either participants or providers due to nature of intervention (receipt of parent aide services or case management as usual). However, data collectors were blinded to participant allocation. **Did participants reflect target group?** Partly. Some key groups were not eligible to participate: parents under 18, parents with	**Was the exposure to the intervention and comparison as intended?** Yes. Fidelity and treatment integrity were checked by logging services offered to families in intervention and control conditions. Services were tracked across 5 domains - control group should receive only domain 1 (case management) whereas intervention should receive services in domains 1–5. Fidelity check found that, as expected dosage of services was higher (statistically significant) in intervention compared to control as measured by: - average number of contacts - total length of time - total number of services - types of services delivered. **Was contamination acceptably low?** Yes. See information on fidelity and treatment integrity. There was some 'leakage' in terms of delivery of some parent skills guidance in the control condition, but this is described as 'mild'. **Did either group receive additional interventions or have services provided in a different**	**Does the study's research question match the review question?** Yes. The aim of the study is to '… examine the benefits of a home-based paraprofessional parent aide services in reducing physical abuse and neglect in high-risk parents' (p566). **Has the study dealt appropriately with any ethical concerns?** Partly. No mention of receiving ethical approval. However, ethical considerations have been taken in to account, including: - 'Control group' receives delayed treatment, to reflect the fact that they at high risk for abuse and neglect - Informed consent from all participants. **Were service users involved in the study?** Yes. **Is there a clear focus on the guideline topic?** Yes. Study relates to early help for families at risk of physical abuse and neglect. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population is families and	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Not awarded ++ as does not have a UK focus. **Overall validity score:** + Good external validity. Key limitations of the study: - Small effective sample size (n=101) leading to limited statistical power - Insufficiently long follow-up time - although this was determined for ethical reasons - Some key measures only taken at follow-up (e.g. household inadequacy score). – Relatively high drop-out rate.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
psychotic mental illness, parents with substance misuse problems for which they were not actively receiving treatment, parents with an IQ below 60. It is questionable whether it is valid to exclude all of these groups from intervention. No rationale given in the text - presumably these groups are considered less able to engage with the intervention. **Were all participants accounted for at study conclusion?** Partly. 73.2% of participants were traced at follow-up, indicating a drop-out rate of 26.8%, somewhat higher than would be desirable. The proportion of participants not traceable at follow-up was slightly higher in the control compared to intervention: 22% of intervention group were not traceable and 33% of the control group.	**manner?** Partly. Some control parents also received a small amount of parental skill guidance, which was additional to what was intended for the control condition. However, this would serve to reduce rather than inflate estimates of the effectiveness of the intervention. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Yes. Outcome measures were reliable in that: - a combination of self-report and observation measures were used - all measures are recognised measures with established reliability and validity. **Were all outcome measurements complete?** Partly. Some measures were only undertaken at follow up, e.g. the household inadequacy score. This meant that it was not possible to discern trajectories of improvement or deterioration in the intervention or control groups. **Were all important outcomes assessed?** Partly. The study did	caregivers of children and young people aged <18 who are at risk of abuse and neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Setting is in family home. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to early help. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Study considers maltreatment outcomes - incidence and risk of abuse and neglect. **(For views questions) Are the views and experiences reported relevant to the guideline?** Not applicable (not views question). **Does the study have a UK perspective?** No.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	not look at Child Protective Services reports/cases of substantiated maltreatment. The stated reason is because this measure was likely to show surveillance bias - parents in intervention condition significantly more likely to be reported to due to contact with services. **Were there similar follow-up times in exposure and comparison groups?** Yes. Both intervention and comparison assessed at 6 months. **Was follow-up time meaningful?** Partly. Investigators chose 6-month follow-up, partly due to ethical reasons as it was thought that the control group should not wait more than 6 months before receiving full services. However, relatively short follow-up (ideal would have been one year) is cited as a limitation in the discussion section of the study. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. Exposure and comparison groups showed no statistically significant differences in terms of demographic variables or baseline risk		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	scores. **Was intention to treat (ITT) analysis conducted?** No. No ITT analysis conducted, although comparison between follow-up groups was conducted to check for selective attrition. No evidence of this (although note that attrition rate higher in control group). **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. No specific power calculation conducted. However, study itself acknowledges that sample size of participating mothers is low (n=101), which therefore limits statistical power. **Were the estimates of effect size given or calculable?** Partly. Effect sizes using Cohen's d are reported, but not 95% confidence intervals. Between-groups differences are analysed by looking at difference in d scores. But statistically significant differences were calculated using 2-factor ANOVA. **Were the analytical methods appropriate?** Yes. Conversion of data to standardised d scores, and analysis of within- and between-		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	groups effects. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes - p scores given in association with two-factor ANOVA. **Do conclusions match findings?** Partly. The statistical analysis shows no significant differences between intervention and control groups on any indicators, except for higher inadequacy of households in the intervention group, suggesting that this group had deteriorated on this measure to a greater extent than the control group. The reasons for this are not explored in the discussion section, which is disappointing. Statistically significant improvements within the intervention group (and not the control group) were observed for: - Psychological aggression - Parenting stress - Maternal anxiety, and - Parental mastery. However, the extent of improvement did not show a significant difference between the intervention and control groups. This is described in the conclusion as 'promising trends suggesting		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	the benefit of parent aide services'. This seems to be a slightly over-optimistic interpretation of the results, which perhaps would be better described as inconclusive.		

10. Lam WKK, Fals-Stewart W, Kelley ML (2009) Parent training with behavioral couples therapy for fathers' alcohol abuse: effects on substance use, parental relationship, parenting, and CPS involvement. Child Maltreatment 14: 243–54

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: RCT. **Description of theoretical approach?** Partly. The authors do not present a clear hypothesis or theory of change but they note that knowledge regarding the links between parental substance abuse and maltreatment tends to focus on mothers despite evidence showing that fathers are more likely to perpetrate serious abuse, especially where substance abuse is a factor. They go on to discuss the evidence showing the benefits which behavioural couples therapy (for substance use issues) can have for children and suggest that the skills which parents learn in these sessions such as communication and problem-solving 'spillover' into wider family relationships. (pp244–5). When discussing the treatment	**Was the exposure to the intervention and comparison as intended?** Partly. The authors report that session attendance rates suggest that appropriate dosages were delivered. Attendance in the Parent Skills with Behavioral Couples Therapy group was 84%; for the Behavioral Couples Therapy group the rate was 86%; and the rate in the Individual-Based Treatment group was 83%. **Was contamination acceptably low?** Yes. Therapists who delivered the sessions participated in videotaped training sessions in which their adherence to manual guidelines was reviewed. This was assessed using scales developed for earlier Behavioral Couples Therapy trials. Adherence scores ranged between 8.84 to 8.89 across each treatment condition	**Does the study's research question match the review question?** Yes. The study aimed to examine the '... effects of Parent Skills with Behavioral Couples Therapy (PSBCT) on substance use, parenting, and relationship conflict among fathers with alcohol use disorders' (p243). **Has the study dealt appropriately with any ethical concerns?** No. No details on informed consent or ethical approval are given. **Were service users involved in the study?** No. There is no indication that service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. Study is relevant to the early help section of the guideline.	**Overall assessment of internal validity:** + The study appears to be well designed however it is unclear why involvement with Child Protective Services was measured via parental self-report. In addition, the impact which the individual cognitive behavioural therapy sessions had on outcomes is unclear and the sample size is very small. **Overall assessment of external validity:** + Only awarded a + as the study was conducted in the USA. **Overall validity score:** + As the study was conducted in the USA and has some methodological limitations (i.e. a pilot study

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
condition they also note that the parent skills training component is developed from a programme which is a method of improving parenting and child functioning (Forehand et al. 1981; Jones et al. 2005). **How was selection bias minimised?** Randomised. No details are provided on randomisation method. **Was the allocation method followed?** Partly - No details on allocation concealment are provided. **Is blinding an issue in this study?** Part blinding - It would not have been possible to blind participants of providers. Blinding of interviewers would have been possible but there is no indication that this was the case. **Did participants reflect target group?** – Yes, 51 males voluntarily entering outpatient treatment for alcohol abuse were screened. Of these 15 were found to be ineligible. Of the 36 participants eligible for the study, 6 patients or their partners refused to take part leaving a final sample of 30. The	(scores of 7 or above are deemed acceptable.) The authors also report that pairwise contrasts did not detect any significant differences by treatment condition. **Did either group receive additional interventions or have services provided in a different manner?** Partly. Twelve of the 24 weekly sessions in each treatment condition were sessions of Cognitive Behavioural Therapy for alcohol treatment delivered individually to male participants. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Partly. The majority of measures used were reliable and valid however they were all collected via maternal, paternal or child self-report at interview. It should be noted that child maltreatment was measured via parental self-report at interview. It seems feasible that this could have been collected via administrative data - the authors do not explain why this method was chosen.	**Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to early help. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Study measures substance use, parenting behaviours, interparental conflict and violence and child maltreatment. **(For views questions) Are the views and experiences reported relevant to the guideline?** Not applicable (not views question). **Does the study have a UK perspective?** No - USA.	with a very small sample size) it is not possible to award a higher overall validity score.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
study excluded couples in which the female also met Diagnostic and Diagnostic and Statistical Manual of Mental Disorders (4th edn; DSM-IV; American Psychiatric Association 1994) criteria for alcohol abuse or dependence. The study also excluded families in which the couple had been cohabiting for less than 2 years or married for less than 1 year. **Were all participants accounted for at study conclusion?** Yes, 83% of the total sample completed all assessments. Drop-out rates are not reported by group.	**Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. Although it is unclear why official records were not used to measure open Child Protection Services cases rather than parental reports. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Yes. Baseline assessments were conducted within 1 week of admission to the treatment program. Three follow-up assessments were conducted: - at treatment completion; at 6 months; and at 12 months. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. No significant differences were found between the three treatment conditions on sociodemographic or background characteristics (all ps>0.30). **Was intention to treat (ITT) analysis conducted?** Not reported. No indication that ITT analysis		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	was conducted. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. Power calculations and expected effect sizes are not presented. The study is a pilot and as such the sample size is very small. **Were the estimates of effect size given or calculable?** Yes. Effect sizes using r are provided (including Cohen's characterisations: small r=.10, medium r=.30, large r =.50). Clinical significance was defined as r≥.20 **Were the analytical methods appropriate?** Yes. Growth curve modelling within a linear mixed effects model framework. Pairwise comparisons between the experimental condition (Parent Skills with Behavioral Couples Therapy) and the two comparators. Pairwise comparisons between the 2 comparators are not reported. Missing data was addressed via imputation. **Was the precision of intervention effects given or calculable?**		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Were they meaningful? Not reported - p values and confidence intervals are not provided. **Do conclusions match findings?** Yes. Although with regards to the measure of child maltreatment the authors emphasise clinically significant effects for the experimental condition r ≥.20 when this is a small to medium effect in statistical terms.		

11. LeCroy CW and Krysik J (2011) Randomized trial of the healthy families Arizona home visiting program. Children and Youth Services Review 33: 1761–6

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: RCT. **Description of theoretical approach?** Partly. The authors do not present a clear hypothesis or theory of change although they do briefly discuss why certain outcomes are important and can contribute to prevention of child abuse and neglect. **How was selection bias minimised?** Randomised. No details on randomisation process are provided.	**Was the exposure to the intervention and comparison as intended?** Not reported. **Was contamination acceptably low?** Not reported. **Did either group receive additional interventions or have services provided in a different manner?** Partly. It is not clear whether the control group received additional interventions as the authors note that '... control group families were offered opportunities to access services if desired' (p1766).	**Does the study's research question match the review question?** Yes. The study aimed to '... examine the effectiveness of home visiting as a means of improving parental, child, and maternal outcomes and preventing child abuse and neglect' (p1761). **Has the study dealt appropriately with any ethical concerns?** Partly. Participants agreed to take part in the programme (although no information is given on consent to take part in the study). The study protocol was approved by the Institutional Review Board.	**Overall assessment of internal validity:** + The decision to use a number of non-validated scales and a maternal depression measure which proved to be unreliable (all of which were collected via parental report), the failure to measure some outcomes at all assessment points, as well as a lack of methodological detail on issues such as the randomisation process suggest lower internal validity. **Overall assessment of external validity:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Was the allocation method followed? Not reported. Allocation concealment is not reported. **Is blinding an issue in this study?** Part blinding. Although the authors report that home visitors were, as far as was possible, kept unaware of group allocation it seems unlikely that this would have been possible given the nature of the intervention. Similarly, it seems unlikely that participants could have been blinded. Data was collected via interview and it would have been possible for investigators to be blinded although it is not clear if this was the case as no information on the interview process is provided. **Did participants reflect target group?** Yes. The programme aims to work with families at the prenatal or new parent stage - 405 families were screened, of which 372 were randomised to the programme. The authors do not report specifically on how many families agreed to participate in the study but 195 were enrolled in total. Eligibility was first assessed using a screening tool (15-item risk criteria, e.g. teenage mother).	**Were outcomes relevant?** Yes. Outcomes measured included 'violent behaviour' (family violence and aggressive discipline) and 'parenting attitudes' (e.g. inappropriate expectations). The authors decided not to measure official reports to the authorities due to concerns regarding surveillance bias. **Were outcome measures reliable?** Partly. All outcome measures relied on self-report at interview. In addition, a number of outcome measures were created specifically for this study, or were modified versions of other scales. The validity and reliability of these are unclear and it is also unclear why these were created in some cases, e.g. a scale was created to measure domestic/family violence that was similar to the Conflict Tactics Scale (Straus et al. 1998) but it is not apparent why this scale itself was not used. The authors report that the scale used to measure depression proved unreliable and these scores using this were not interpreted due to low alphas.	**Were service users involved in the study?** No. No indication that service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. Study aims to prevent maltreatment and improve child and parental outcomes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Families screened (using a 15 item checklist, e.g. teen mother) and assessed (using Kempe Family Checklist) for risk. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. The intervention is home visiting. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Relevant to early help. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Outcomes measured included 'violent behaviour' (family violence and	Only given a + as the study was carried out in the USA. **Overall validity score:** + As the study was conducted in the USA and has some internal validity issues (e.g. use of non-validated scales, data collected via parental report, the failure to measure some outcomes at all assessment points) it is not possible to award a higher validity score.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
If the family screened positively (no information on threshold required) a survey was administered. This was a modified version of the Kempe Family Checklist. If the score was higher than 25 for either parent then the programme was offered to the family. No further exclusion criteria are listed. No details on the proportion of African-American participants is provided in demographic data and the majority of participants in both intervention and control groups were Hispanic (64.9% and 54.6% respectively). **Were all participants accounted for at study conclusion?** Yes. Drop-out rates were acceptable and comparable. At 6 months assessment, 94% of the intervention group and 91% of the control group completed assessments. At the 1 year assessment, retention rates were 88% in the intervention group and 89% in the control group.	**Were all outcome measurements complete?** Partly. Some measurements do not appear to have been completed at certain points, e.g. maternal engagement in school or training at baseline or emotional loneliness at the one year assessment. **Were all important outcomes assessed?** Partly. The authors did not use official child protection services as a measurement of incidence of abuse. They state that these are not recommended as valid outcome measures and are likely to show surveillance bias. In addition, the lack of child outcomes (e.g. development) seems problematic given the stated aims of the study and programme. **Were there similar follow-up times in exposure and comparison groups?** Yes. Both the intervention and control groups were assessed at 6 months of age and 1 year of age. However, there is little detail provided regarding the point at which screening and enrolment took place and given the fact that the authors state that the programme is targeted at new parents or those in the prenatal stage	aggressive discipline) and 'parenting attitudes' (e.g. inappropriate expectations). **(For views questions) Are the views and experiences reported relevant to the guideline?** Not applicable (not views question). **Does the study have a UK perspective?** No. Conducted in the USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	there may have been some variation in the length of time which the families had been involved in the programme. Ages of children are not provided in the demographic data. **Was follow-up time meaningful?** Partly. Assessment at 6 months and one year of age seems acceptable (resulting in low drop-out rates). The authors note that one year is a useful assessment point for certain measures e.g. the Parent -Child Conflict Tactics scale as this '... is the period when parents are apt to react more punitively to their infants' (p1763). However, this time period means that the long-term impacts of the programme could not be measured. The authors themselves note this and make specific comments with regards to certain outcomes such as abusive behaviours which can increase during the toddler period (e.g. spanking). NB. The authors originally intended to collect data over a 5-year period, but funding was eliminated due to the financial climate. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** No.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	The authors report that the control and intervention group were similar on most characteristics at baseline although there were statistically significant differences found in mother's average age, receipt of prenatal care, health insurance, employment of mother, ownership of a car. The authors also report that involvement with child protective services as a parent was a statistically significant difference, with more participants in the intervention group reporting involvement. As a result the authors decided to use an ANCOVA model to covary any differences. **Was intention to treat (ITT) analysis conducted?** Yes. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. Power calculations and expected effect sizes are not provided. The sample size seems reasonable although the authors note that it was 'somewhat underpowered' (p1763) and chose to define statistical significance at p=<.10 as a result. **Were the estimates of effect**		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	size given or calculable? Not reported. Effect sizes are not provided. **Were the analytical methods appropriate?** Yes. The authors used an ANCOVA model to allow variables that were significantly different at baseline to be used as covariates. 'The analysis used a series of a priori contrasts' (p1763). The dependent measures were considered 'conceptually independent' and were treated as an independent test. 'However, measures that were highly correlated with other measures were not included' (p1763). **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes - p values are provided. **Do conclusions match findings?** Yes. Although the text is not always clear that some measures were not assessed at baseline, e.g. maternal engagement in school or training.		

12. Mejdoubi J, van den Heijkant SCCM, van Leerdam FJM et al. (2015) The effect of VoorZorg, the Dutch nurse-family partnership, on child maltreatment and development: a randomized controlled trial. PloS one 10: e0120182

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: To investigate the effectiveness of VoorZorg, a Dutch adaptation of the Nurse-Family Partnership (NFP) in preventing child maltreatment. **Description of theoretical approach?** Partly. No description of why NFP, upon which ZV is based is effective, but does state that NFP has been found to be effective in 3 RCTs conducted in the US, including the Elmira trial which showed at ages 2 and 15 NFP children were less likely to have CPS reports. States that this is the first trial to conduct an RCT of NFP outside the US. (p 2). **How was selection bias minimised?** Randomised. **Was the allocation method followed?** Yes. **Is blinding an issue in this study?** Blinding. Single blinding - researchers as it was obvious to participants which group they were in. No way of avoiding this when comparing with usual care. **Did participants reflect target group?** Yes. Although fewer	**Was the exposure to the intervention and comparison as intended?** Partly. High rate of attrition from baseline to 24 months follow up Control: n=223 to n=93 for IT-HOME and CBCL outcomes, n=223 to n=164 for 36 month CPS outcome Intervention n=237 to n=130 for IT-HOME and CBCL outcomes, n=237 to n=168 for 36 month outcomes. **Was contamination acceptably low?** Not reported. **Did either group receive additional interventions or have services provided in a different manner?** Not reported. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Yes. IT-HOME and CBCL are validated tools CPS reports accepted as measure of risk/incidence of abuse or neglect. **Were all outcome measurements complete?** Partly. Smaller	**Does the study's research question match the review question?** Yes. Looks at effectiveness of programme in reducing child maltreatment, measured by incidence of CPS reports 36 months after birth. **Has the study dealt appropriately with any ethical concerns?** Yes. Written consent obtained from participants and agreement from them to obtain CPS data on their children. Written forms scanned and stored in digital archive. Ethics approved by Medical Ethical Committee of VU University Medical Centre. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes - relates to 4.3 c) targeted activities to prevent child abuse. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Yes relates to both children at risk of abuse or neglect and mothers of those children.	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** ++ Only country is not exactly relevant. **Overall validity rating:** + Relatively high attrition rate for the study (32.8% for whole study sample).

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
women were available to be analysed for CPS reports than other outcomes as only 8 CPS regions gave their permission. **Were all participants accounted for at study conclusion?** Partly. Attrition rate at 24 months in intervention group was 25% and in control group was 41%. Missing data were imputed as part of intent to treat analysis.	groups for both control and intervention were measured for incidence of abuse/neglect **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Yes. 6 months, 18 month and 24 months for IT-HOME and CBCL measurements, 36 months for CPS reports. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. **Was intention to treat (ITT) analysis conducted?** Yes. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Partly. Power calculation was conducted for outcome measured by different study of same trial - smoking behaviour, where power calculation was 0.8, alpha 0.05. **Were the estimates of effect size given or calculable?** Yes. **Were the analytical methods appropriate?** Yes. Poisson regression models to assess difference between groups for CPS reports	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Children's own homes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. 4.3c. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Incidence of abuse or neglect; quality of parenting; child's wellbeing. **Does the study have a UK perspective?** No. Dutch.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	an CBCL/1.5 to 5 scores Moderation analyses conducted on primary outcome (CPS reports) for ethnicity and gender of child. Multiple Imputation analyses conducted on missing CBCL data at 24 months, validated by sensitivity analyses - 50 imputed datasets generated as recommended Multiple linear regression first used to assess difference in IT-HOME scores with a mixed model analysis conducted to look at longitudinal difference- MI not conducted due to higher power of mixed model analyses. All analyses adjusted for confounders and effect modifiers- region; age; ethnicity; gender of child; age of mother; weeks of gestation and birth weight). Attrition analysis also conducted. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. **Do conclusions match findings?** Yes.		

13. Nelson HD, Selph F, Bougatsos C et al. (2013) Behavioral Interventions and Counseling to Prevent Child Abuse and Neglect: Systematic Review to Update the U.S. Preventive Services Task Force Recommendation. Rockville: Agency for Healthcare Research and Quality

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. 'This systematic review addresses the effectiveness and adverse effects of behavioral interventions and counseling to prevent child abuse and neglect for children at potentially increased risk. This review focuses on children without obvious signs or symptoms of abuse or neglect who are seen in health care settings' (p1). **Adequate description of methodology?** Yes. 'An investigator abstracted data about the study design and setting, participant characteristics, data collection procedures, numbers enrolled and lost to follow-up, methods of exposure and outcome ascertainment, analytic methods including adjustment for confounders, and outcomes. A second investigator confirmed the accuracy of data' (p7). Consistency judged according to whether outcomes generally in same direction of effect, and range of effect sizes were narrow. Meta-analysis not possible due to	**Inclusion of relevant individual studies?** Yes. **Study quality assessed and reported?** Yes - using predefined criteria developed by the USPSTF (US Preventive Services Task Force Quality Rating Criteria), two investigators rated the quality of studies (good, fair, poor) and resolved discrepancies by consensus (Appendix A5). **Do conclusions match findings?** Yes. Balanced conclusions with relation to findings.	**Does the study's research question match the review question?** Yes. 'This systematic review addresses the effectiveness and adverse effects of behavioral interventions and counseling to prevent child abuse and neglect for children at potentially increased risk. This review focuses on children without obvious signs or symptoms of abuse or neglect who are seen in health care settings' (p1). Comparative studies (RCTs) but 8/17 included studies published after 2004. However, pre-2004 studies have been included for data extraction. **Has the study dealt appropriately with any ethical concerns?** Not reported - not applicable. **Were service users involved in the study?** Not reported. **Is there a clear focus on the guideline topic?** Yes. 'This sys-	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** + **Overall score:** + Trials were limited by heterogeneity, low adherence, high loss to follow-up, and lack of standardised measures

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
heterogeneity of participants, interventions, outcome measurements, follow-up periods and data analysis. The review was reviewed by content experts, USPSTF members, AHRQ Project Officers, and collaborative partners. Detailed presentation of quantitative data from each study when data available (Table 4-9). **Rigorous literature search?** Yes. Detailed search strategy (Appendix A1); 'MEDLINE and PsycINFO (January 2002 to June 2012), Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews (second quarter 2012), Scopus, and reference lists were searched for English-language trials' (Abstract).		tematic review addresses the effectiveness and adverse effects of behavioral interventions and counseling to prevent child abuse and neglect for children at potentially increased risk. This review focuses on children without obvious signs or symptoms of abuse or neglect who are seen in health care settings' (p1). **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. 'This review focuses on children without obvious signs or symptoms of abuse or neglect who are seen in health care settings' (p1). **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Homes and health care settings (prenatal clinics; hospitals). **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Various home visiting and parenting programmes. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. CPS	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		involvement, emergency department admission, hospital visits, reports of harsh parenting, etc. **Does the study have a UK perspective?** No. Mainly US, only 1 of 10 trials conducted in UK.	

14. Peacock S, Konrad S, Watson E et al. (2013) Effectiveness of home visiting programs on child outcomes: A systematic review. BMC Public Health 13: 17

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes - clear objectives. **Adequate description of methodology?** Yes. 'A principal reviewer assessed all the papers, and one of two secondary reviewers independently evaluated their relevance, with a third to adjudicate if needed. When necessary, we contacted researchers to clarify components of their research' (p2). Narrative data analyses of studies according to outcomes. **Rigorous literature search?** Yes. Search of 7 databases including CINAHL PLUS, Cochrane Library, ProQuest Dissertations and Theses, EMBASE, MEDLINE, PSYCINFO and Sociological Abstracts databases (pub date 1990 to May 2012); detailed search strategy.	**Inclusion of relevant individual studies?** Yes. **Study quality assessed and reported?** Yes. Based on validity tool (total scores of 15), 3 scores for each on study design/allocation, matched cohort; attrition; control of confounders, measurement tools description and appropriate statistical analyses. 6 RCTs scored 15/15; 15 RCTs scored 13-14/15. Only studies score 13/15 were included. **Do conclusions match findings?** Yes. Conclusions fair and balanced as authors also aware of limitations: access to papers, unsuccessful contact with authors (of studies) for more information; publication bias, selective reporting within studies etc.	**Does the study's research question match the review question?** Yes. Review of 21 studies. 17 studies were included which met the scope of our review, however four did not (originating from Chile, Bangladesh, Jamaica and South Africa) and for the purposes of data extraction have been excluded. **Has the study dealt appropriately with any ethical concerns?** Not reported - not applicable. **Were service users involved in the study?** Not reported. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes - at risk families and mothers. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes - not explicit. Home visitation programmes are likely to be rolled out in homes and primary healthcare settings.	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** + **Overall score:** + Study does not report the outcome measures used in studies where no significant impact was found. The authors report that the 'findings of this review must be considered in light of the potential for publication bias, selective reporting within studies and methodological limitations within the included studies' (p13).

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Does the study relate to at least 1 of the activities covered by the guideline? Yes. Home visiting -targeted activities /interventions to prevent child abuse and neglect. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Child abuse and neglect; report on health assessment: physical growth; up-to-date immunisations; hospitalisations, illness and injuries; report on developmental delays: psycho-motor and cognitive development; child behaviour; language development. **Does the study have a UK perspective?** No. Included studies conducted in different countries, 14 in the US; 2 in Ireland and 1 in UK.	

15. Pereira M, Negrão M, Soares I et al. (2015) Decreasing harsh discipline in mothers at risk for maltreatment: A randomized control trial. Infant Mental Health Journal 35: 60413

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: To assess the effectiveness of the Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD) at improving maternal sensitivity and reducing harsh discipline among 43 severely deprived Portuguese mothers of children aged 1–4 years, for whom there were concerns raised about the caregiving environment. **Description of theoretical approach?** Yes. Links the attachment theory framework which underpins the Ainsworth et al. maternal sensitivity construct (Ainsworthet al. 1978) to Milner's 4-stage social-information-processing model for understanding the risk of the use of harsh parenting. (Milner 1993, 2003). 'The parent's lack of positive attachment-related skills such as sensitivity and empathy for the child leads to negative interpretations of child behavior and potentially harsh discipline practices' (p606). Also draws on insights from coercion theory- whereby parents and children can get into negative interaction cycles, through child-resistance and an escalation to harsh discipline to counteract this, and states that as this pattern can	**Was the exposure to the intervention and comparison as intended?** Yes. **Was contamination acceptably low?** Not reported. **Did either group receive additional interventions or have services provided in a different manner?** No. Those that received usual care of the telephone sessions were signposted to GPs or health agency if they sought explicit help and advice from researchers. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Yes. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** No, 1 month post-test only. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes.	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Yes. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. Reduction in harsh discipline. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Parents at risk of maltreating their children. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Own homes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Early help. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **Does the study have a UK perspective?** No. Portuguese.	**Overall assessment of internal validity** - **Overall assessment of external validity** + Entirely relevant but not UK. **Overall validity score** - Small sample, no blinding, lack of information on whether study was sufficiently powered, also high levels of attrition as study progressed and unclear what impact this had on representativeness of sample.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
start very early in life, preventative efforts are needed. Does not state why VIPP-SD is effective beyond that it has been found to be effective in several other RCTS when used with other samples including adoptive parents, children at risk of externalising problems, children with dermatitis, insecurely attached mothers with temperamentally reactive infants, mothers with postnatal eating disorders. This is the first RCT to test the method with disadvantaged mothers and in order to specifically reduce harsh discipline. **How was selection bias minimised?** Randomised. **Was the allocation method followed?** Yes. **Is blinding an issue in this study?** No blinding. **Did participants reflect target group?** Partly. Sample reflected the kinds of mothers, but target selection sample was 132, and the end sample was 44, so there was a high attrition rate from initial recruitment. Then further data collection issues further reduced the analysed sample to 43.	**Was intention to treat (ITT) analysis conducted?** No. Only participants with complete set of data analysed which saw n=44 sample reduced to n=43. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. **Were the estimates of effect size given or calculable?** Yes. **Were the analytical methods appropriate?** Yes. Outliers adjusted for. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. p values but not confidence intervals. **Do conclusions match findings?** Yes.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Were all participants accounted for at study conclusion? Yes. All families accounted for. However, there were high levels of dropout following the pretest stage (40%), and relatively high levels of drop-out during the main intervention (20%).			

16. Robling M, Bekkers M-J, Bell K et al. (2015) Effectiveness of a nurse-led intensive home-visitation programme for first-time teenage mothers (Building Blocks): a pragmatic randomised controlled trial. Lancet: 1–10

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: To investigate the effectiveness of a nurse-led home visiting programme for first-time mothers aged 19 years or younger in a UK context 24 months after birth. The primary aims of the intervention were to reduce tobacco use by the mother, proportion of subsequent pregnancies within the 24 months, birthweight of the baby, and number of emergency department attendances and hospital admissions. We have also extracted data collected by the study which relates to outcomes of interest to this review. **Description of theoretical approach?** Yes.	**Was the exposure to the intervention and comparison as intended?** Yes. **Was contamination acceptably low?** Not reported. **Did either group receive additional interventions or have services provided in a different manner?** No. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Partly. Query measures used to assess safeguarding outcomes. These are:	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Yes. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Young first-time mothers.	**Overall assessment of internal validity** + Bias of using only risk factor of age of mother. Different groups of sample used for different outcomes. **Overall assessment of external validity** + Due to mismatch in primary outcomes. **Overall validity score** + This study has been rated as moderate due to the rigorous design and detailed reporting of

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
How was selection bias minimised? Randomised. Randomisation stratified by site and minimised by gestation, smoking and preferred language of data collection. **Was the allocation method followed?** Yes. **Is blinding an issue in this study?** Part blinding. Mothers and field-based researchers (who could also have been assessors) not masked to group allocation but assessors conducting computer-assisted telephone interview were masked. **Did participants reflect target group?** Yes. **Were all participants accounted for at study conclusion?** Partly. Full evaluation report reports voluntary withdrawals from the programme, and mandatory withdrawals (for example, due to miscarriage or stillbirth). The rates of mandatory and voluntary withdrawals were relatively low (5.0% and 6.7% respectively). However, rates of incomplete assessment data were relatively high. For example, 18% of the sample did not	- safeguarding processes as reported in GP notes – query re. how comprehensive this would be. - would all safeguarding processes always be captured? - self-reported referrals to social services - query regarding how accurate self-report likely to be - self-reported levels of intimate partner violence. Suggest that a more accurate measure of safeguarding concerns would have been to look at children's social care records. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Partly. Different follow-up times for different outcomes **Was follow-up time meaningful?** Yes. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. **Was intention to treat (ITT) analysis conducted?** Yes. **Was the study sufficiently powered to detect an intervention**	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Own homes of mothers and children. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Early help. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Partly. The primary outcomes of interest for this study had more of a health focus: tobacco use, birthweight of baby, proportion of second pregnancies within 24 months, emergency attendance and hospital admissions within 24 months. The study does look at safeguarding and referrals to social services as two of its secondary outcomes, however these are as reported in GP records and self-reported respectively. The study has not used data from children's social care directly. **Does the study have a UK perspective?** Yes.	data. However, we considered rating this as poor for 2 reasons: 1) Possible risk of bias arising from high levels of missing data. In line with the Cochrane guidelines we have considered whether the missing data presents a risk of bias in terms of 'the amount and distribution across intervention groups, the reasons for outcomes being missing, the likely difference in outcome between participants with and without data, what study authors have done to address the problem in their reported analyses, and the clinical context' (http://handbook.cochrane.org/). For the safeguarding data provided by GPs our view is that the risk of bias resulting from attrition is low because rates of returned information from GPs is similar for both intervention and control. For self-report data we suggest that there may be some risk of bias due to higher levels of failure to complete assessments at 24 months amongst the control group (21.9% incomplete) compared to the intervention group (14.3% incomplete). The reviewing team have calculated this difference as being significant (chi-square=13.6, $p<0.01$).

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
complete the 24 month assessment data. For example: Safeguarding (data obtained from GP) - collected data represent only 57.4% of randomised participants. Referrals to social services (data obtained via self-report at 24 months) - collected data represent 68.1% of randomised participants. In line with the Cochrane guidelines, we have considered whether the missing data are likely to bias the findings. For the safeguarding data provided by GPs our view is that the risk of bias resulting from attrition is low because rates of returned information from GPs is the same for both intervention and control. For self-report data we suggest that there may be some risk of bias due to higher levels of failure to complete assessments at 24 months amongst the control group (21.9% incomplete) compared to the intervention group (14.3% incomplete). The reviewing team have calculated this difference as being significant (chi-square=13.6, p<0.01).	**effect (if one exists)?** Not reported. **Were the estimates of effect size given or calculable?** Yes. **Were the analytical methods appropriate?** Yes. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. **Do conclusions match findings?** Yes.		2) The validity of the study outcome measures in terms of our principal outcome of interest - incidence of abuse and neglect. Safeguarding concerns are measured in the study via GP records and self-report. We are concerned that both of these methods are likely to under-report safeguarding issues, and consider that a better outcome measure would have been to use children's social care data. However, the problem of under-reporting should be equally present in both intervention and control groups. Based on these considerations we decided to retain a rating of 'moderate'.

17. Sanders MR, Pidgeon AM, Gravestock F et al. (2004) Does parental attributional retraining and anger management enhance the effects of the Triple P-Positive Parenting Program with parents at risk of child maltreatment? Behavior Therapy 35: 513–35

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: RCT. **Description of theoretical approach?** Partly. The authors do not present a logic model or comprehensive theory of change but they note that evidence shows that there may be a relationship between dysfunctional parental attributions and negative child outcomes. They go on to suggest that interventions targeted at parents who maltreat may be more effective where they include a cognitive behavioural component. (pp515–16). **How was selection bias minimised?** Randomised. No details on the randomisation process are given. **Was the allocation method followed?** Not reported. Allocation concealment is not reported. **Is blinding an issue in this study?** Part blinding. Observational measures were recorded by video and interactions were coded by raters blind to group assignment and stage of assessment. It would	**Was the exposure to the intervention and comparison as intended?** Yes. Although 12% of the total sample failed to complete the treatment programme. Treatment sessions were delivered using written protocols and adherence to these was recorded by co-facilitators using checklists. These were analysed to ensure that the required content had been covered. Interrater reliability between coder and provider showed a high level of agreement. In addition, a research assistant randomly viewed video recordings of 15 sessions from each condition and coded these according to protocol adherence. Adherence levels (mean) were high for each condition. **Was contamination acceptably low?** Partly. The authors report that participants in the standard condition may have been incidentally exposed to attributional training as a result of within group socialisation processes. **Did either group receive additional interventions or have services provided in a different**	**Does the study's research question match the review question?** Yes. The study aimed to compare the effectiveness of '... an enhanced group behavioral family intervention (EBFI) for parents at risk of child maltreatment ..." which addressed parental negative attributions and anger management to a "... standard-care group parent training intervention ...' (p516). **Has the study dealt appropriately with any ethical concerns?** No. No information regarding informed consent or approval of the study protocol is provided. The authors do note that their decision not to use a no-treatment or wait-list control group was due to ethical concerns regarding withholding treatment from participants at risk of child abuse. **Were service users involved in the study?** No. No indication that service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. Relevant to early help section of the guideline.	**Overall assessment of internal validity:** + On the whole the study seems well designed although there is a lack of methodological detail and the follow up period was relatively short. In addition, a small sample size is a significant limitation. **Overall assessment of external validity:** + Only awarded a + as the study was conducted in Australia. **Overall validity score:** + As the study was conducted in Australia and there are some internal validity concerns (e.g. small sample size and short follow-up) it is not possible to award a higher overall validity score.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
not have been possible to blind providers or participants. No details are provided on whether baseline or follow-up interviews were conducted by investigators who were blinded to group assignment although this would have been possible. **Did participants reflect target group?** Yes. No details on the number of families screened for eligibility, the number ineligible or participation rates of eligible families are provided. Selection criteria: At least 1 notification to Families, Youth and Community Care Queensland (did not need to be substantiated) and/or parental concerns regarding anger management in relation to the child, and an elevated score on 3 subscales of the State-Trait Anger Expression Inventory (STAXI, Spielberger 1996). Families were excluded if they were:- currently receiving family therapy or psychotherapeutic interventions in relation to child behaviour; a child or parent had significant intellectual impairments. No families were excluded on these criteria. The authors also report that the study participants were families with a child between the ages of two to	**manner?** No. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Partly. All measures were made using established and validated scales although all appear to rely on parental self-report. It may also be important to note that the authors report that the some measures (e.g. the Parental Anger Inventory - used as a measure of risk of maltreatment) only showed moderate reliability, internal consistency, and correlation with other measures. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Partly. Assessments took place pre-treatment, post-treatment and at the six month point. Six months is quite a short follow-up time and	**Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Parents '... experiencing significant difficulties in managing their own anger in their interactions with their preschool-aged children ...' (p513). **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Relates to early help. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. The study measures risk of maltreatment, parenting and parental adjustment, and child behaviour. **(For views questions) Are the views and experiences reported relevant to the guideline?** Not applicable (not views question). **Does the study have a UK perspective?** No. The study was conducted in Australia.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
seven but it is not clear if this is an eligibility criterion. The study's outreach strategy is reported as '... parents who were concerned about their anger or that they would harm their child rather than concerns specifically about child behavioral problems' (p518). The majority of participants were female. Ethnic background is not reported. Only a small percentage of participants had been referred to authorities for allegations of abuse or neglect (did not need to be substantiated) which was one of the two main eligibility criteria (SBFI 4%; EBFI 6%). **Were all participants accounted for at study conclusion?** Yes. In the total sample (n=98), drop-out rates were acceptable (12%) but these differed by group, with 8 families from the experimental condition dropping out and 4 from the standard condition. Of the 86 families who completed their treatment programme, 2 families (2%) could not be contacted at post-intervention assessment point and a further 2 (2%) could not be reached at the 6-month assessment point. These families lost to follow-up are not reported by group.	long-term impacts of the programme would not be apparent at this stage. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. They authors report that the groups were 'well matched' on family background and demographic characteristics at baseline. The authors also used ANOVA to compare the groups on key outcome measures at baseline. No significant differences were found on any measure. **Was intention to treat (ITT) analysis conducted?** Not reported. No indication that ITT analysis was conducted. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. Power calculations and expected effect sizes are not reported. The total sample size was small (n=98). **Were the estimates of effect size given or calculable?** Not reported. Effect sizes and confidence intervals are not provided.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Were the analytical methods appropriate? Yes. The analytical methods seem appropriate: 2 x 3 MANOVA. Clinically reliable changes were also assessed. Due to large number of measures, use of conservative alpha level of .01 across all analyses to control for the potential inflation of Type 1 error. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes - p values are provided. **Do conclusions match findings?** Yes. Although the authors' discussion of 'long-term intervention effects' is based on data collected at six months. (p528).		

18. Scudder AT, McNeil CB, Chengappa K et al. (2014) Evaluation of an existing parenting class within a women's state correctional facility and a parenting class modeled from parent–child interaction therapy. Children and Youth Services Review 47: 238–47

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: RCT. **Description of theoretical approach?** Partly. The authors provide a review of the evidence on Parent Child Interaction Therapy and a history of the development	**Was the exposure to the intervention and comparison as intended?** Not reported. No detail is provided on exposure. It may be important to note that the intervention was modified to enable delivery in a correctional environment	**Does the study's research question match the review question?** Yes. The aim of the study is to determine whether parenting classes devised using Parent Child Interaction Therapy are more effective than a traditional	**Overall assessment of internal validity:** + This is a well-designed study however there are some limitations including a small sample size, short-term follow-up, and collection of

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
of the programme citing Hanf's 2-stage model (Reitman and McMahon 2012) and Baumrind's (1967) theory regarding parental nurturance but they do not outline a theory of change. **How was selection bias minimised?** Randomised. Carried out using an online randomiser. **Was the allocation method followed?** Yes. Although allocation concealment is not reported (and the authors state that a clerical worker was involved with the randomisation process), testing of between-group differences at pre-treatment showed that there were no significant differences. **Is blinding an issue in this study?** Part blinding. Investigators (those who coded interactions) were blinded to treatment condition but it would not have been possible to blind participants or providers. **Did participants reflect target group?** Yes. Women were eligible if they had been recommended for a parenting programme by facility staff at intake. Women not fluent in English were excluded. At least	(e.g. in conventional PCIT progression usually depends on 'mastery' of certain skills. In this program this was not required for progression or graduation.) **Was contamination acceptably low?** Not reported. No details provided on contamination. **Did either group receive additional interventions or have services provided in a different manner?** No. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Partly. All outcomes were measured using validated scales however some are self-reported and others were determined via observation in a role-play situation (i.e. with another researcher). **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Partly. Child outcomes were not measured, which the authors themselves note as a limitation.	parenting programme delivered to incarcerated mothers in enhancing parenting skills and stress, knowledge of child development, and child abuse potential. **Has the study dealt appropriately with any ethical concerns?** Yes. The study protocol was approved by the West Virginia Institutional Review Board. Consent from participants was given and they were aware that they could withdraw at any time without penalty. **Were service users involved in the study?** No. No indication that there was input at the design stage. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes.	data via role-play (i.e. not with the mothers child). **Overall assessment of external validity:** + Only given a + as the study was not conducted in the UK. In addition, the Guideline Committee should bear in mind that the intervention was delivered to incarcerated mothers. **Overall validity score:** + This is a well-designed study however there are some limitations including a small sample size, short-term follow-up, and collection of data via role-play (i.e. not with the mothers child). In addition, as the study was conducted in the USA and was delivered to incarcerated mothers there are limits to the external validity of the study.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
one of the women's children had to be under 12 which the authors state is due to the proven efficacy of Parent Child Interaction Therapy for children under 12. 84 women were selected and only two did not complete pre-treatment assessment (placed back on the waitlist). **Were all participants accounted for at study conclusion?** Yes. The authors report that 71 women completed pre-post measures but this does not seem to match the data in table 1 on p243. There was a total attrition rate of 13%. Three participants dropped out from the intervention group and ten from comparison group (25%) which is quite high.	**Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Partly. Assessments were carried out at pre and post-treatment. Follow-up at post-treatment did not allow the long-term impacts of the programme to be measured. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. Testing of between-group differences at pre-treatment showed that there were no significant differences. **Was intention to treat (ITT) analysis conducted?** Not reported. No indication that ITT analysis was carried out. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. No power calculation or expected effect sizes are provided. The sample size is quite low. **Were the estimates of effect size given or calculable?** Yes.	**(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **(For views questions) Are the views and experiences reported relevant to the guideline?** Not applicable (not views question). **Does the study have a UK perspective?** No. Conducted in the USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Effect sizes using Cohen's d are provided. **Were the analytical methods appropriate?** Partly. Mixed between-within ANOVAs used for all measures. However, only one set of p values reported - unclear whether these are main effect of group, or interaction of group by time. Child abuse potential scores were examined with and without elevated lie scores. **Was the precision of intervention effects given or calculable? Were they meaningful?** No. Confidence intervals and p values are not reported. **Do conclusions match findings?** Yes. On the whole the authors' conclusions do match the findings. The study found that inappropriate expectations decreased to a greater extent in the comparison group which the authors seem to attribute to a lack of emphasis on this issue in the intervention, however it is not clear if this issue is directly addressed in the existing facility programme. The reasons for similar decreases in child abuse potential between the two		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	groups are not discussed.		

19. Silovsky JF, Bard D, Chaffin M et al. (2011) Prevention of child maltreatment in high-risk rural families: A randomized clinical trial with child welfare outcomes. Children and Youth Services Review 33: 1435–44

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: RCT. **Description of theoretical approach?** Yes. Intervention is based on developmental-ecological theory of the etiology of physical child abuse and neglect (e.g. Belsky 1993). **How was selection bias minimised?** Randomised. Randomised using computer algorithm. **Was the allocation method followed?** Yes. Study also reports no significant differences in demographic variables between intervention and control groups. However, no analysis of existing difference in risk factors between intervention and control groups is reported. **Is blinding an issue in this study?** Part blinding. Allocation should have been concealed from assessors - however this is not reported.	**Was the exposure to the intervention and comparison as intended?** Partly. Exposure was as expected in terms of fidelity to both interventions. However, engagement with the control service (home-based mental health service) was significantly poorer than with intervention, with higher attrition rate in the control condition and delivery of significantly fewer hours of service. **Was contamination acceptably low?** Yes. **Did either group receive additional interventions or have services provided in a different manner?** No. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Yes. Standardised measures used across domains.	**Does the study's research question match the review question?** Yes. Study is examining impact of a child maltreatment prevention intervention. **Has the study dealt appropriately with any ethical concerns?** Yes. Human subject protection approval obtained from appropriate Institutional Review Board. Informed consent obtained from all participants. **Were service users involved in the study?** Yes. **Is there a clear focus on the guideline topic?** Yes. Study relates to early help for families at risk of child abuse and neglect. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population is families at risk of child abuse and neglect.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Not given ++ as does not have UK perspective. **Overall validity score:** + Good quality study, key limitations are failure to report concealment of allocation from assessors; lack of analysis of existing risk factors in treatment groups; exclusion of higher risk families (although some studies would suggest that these groups are more amenable to treatment).

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Did participants reflect target group? Partly. Study excluded: 1. Families with current child welfare case involvement 2. Families where primary caretaker has a substantiated report of child sexual abuse 3. Any conditions which would prevent primary caregiver from providing valid self-report data (e.g. severe psychosis, severe learning disability) 4. Families with more than two past reports of CPS involvement. Justification for excluding families with more than two past reports of CPS involvement is that 'typically prognosis for families with a small number of past referrals is better' (p1437). This would suggest that excluding families with >2 previous CPS reports could have the effect of over-estimating effectiveness of intervention. **Were all participants accounted for at study conclusion?** Yes. Intention-to-treat analysis meant that all participants accounted for.	**Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. No statistical difference in timing of post-service and follow-up assessments (measured in days). **Was follow-up time meaningful?** Yes. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Partly. Exposure and comparison groups showed no significance in demographic characteristics. Does not appear to be any statistical analysis of baseline risk scores. **Was intention to treat (ITT) analysis conducted?** Yes. Intention to treat analysis used and missing data imputed. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Yes. Study	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Setting is participants' homes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to early help. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Study measures incidence and risk of child abuse and neglect. **Does the study have a UK perspective?** No.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	reports that: The obtained sample size was powered to detect a minimum 2-year survival hazard ratio of 0.212. The minimum detectable effect size for the generalised linear mixed modelling hypothesis tests ranged from 0.35 to 0.62. These are within acceptable limits, although it is questionable whether these were the appropriate statistical techniques. **Were the estimates of effect size given or calculable?** Yes. No effect sizes reported, but are calculable based on data in paper. **Were the analytical methods appropriate?** Partly. The study uses statistical modelling rather than significance testing to estimate differences within and between groups. The decision to use survival analysis, rather than, for example, a binomial measure of future Child Protective Services referral is not fully explained. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes - p values for test statistics provided.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Do conclusions match findings? Yes. Conclusions are clear about significant and non-significant results.		

20. Stover C (2015) Fathers for Change for Substance Use and Intimate Partner Violence: Initial Community Pilot. Family Process 54:600–9

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim: The assess the feasibility and efficacy of the Fathers for Change programme for men with history of intimate partner violence (IPV) and substance abuse, compared to individual drug counselling. **Description of theoretical approach?** Yes. Based on the evidence that shows greater incidence of IPV in co-parenting relationships of opioid-dependent men; that fathers with co-occurring IPV and substance abuse have more negative co-parenting relationships, and that negative co-parenting can mediate the relationship between having a father with substance abuse, witnessing IPV and child behavioural problems. Positive co-parenting inversely has been found to be pro-	**Was the exposure to the intervention and comparison as intended?** Yes. **Was contamination acceptably low?** Not reported. **Did either group receive additional interventions or have services provided in a different manner?** No. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Yes. Use of validated measures. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes.	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Partly. Ethics of paying participants is not discussed (p603). **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. Preventing child maltreatment through reducing risk factors of substance abuse and reducing risk of witnessing IPV, increasing parenting quality **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Parents at risk of/inflicting child maltreatment	**Overall assessment of internal validity** - Mentions not all participants finished treatment (67% intervention vs. 33% control) but not how this affected the analysis or provides break down of attrition rates. Does not comment on effect size, just significance. Small sample but pilot study. **Overall assessment of external validity** + US focus, lack of clarity on setting. **Overall validity score** - Internal validity issues prevent higher rating

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
tective and result in better adjustment. Design of intervention based on evidence that there can be a reduction in the transmission of IPV through generations if the father perpetrators are supported to recognise the impact of their violence (Guille 2004), that there is a significant subset of men who perpetrate IPV who are concerned about the impact of IPV on their children and that this concern can be a motivating factor for seeking treatment. It is also based on evidence that integration of programmes tackling both IPV and substance abuse can be effective. (p601). **How was selection bias minimised?** Randomised. **Was the allocation method followed?** Yes. **Is blinding an issue in this study?** Blinding. Research assistants blinded to group participation status throughout and at 3 month follow up **Did participants reflect target group?** Yes. **Were all participants accounted for at study conclusion?** Yes.	**Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** No. 3 months quite short follow up time **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. **Was intention to treat (ITT) analysis conducted?** Not reported. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. Power calculation not given but small sample. **Were the estimates of effect size given or calculable?** No. **Were the analytical methods appropriate?** Yes. **Was the precision of intervention effects given or calculable? Were they meaningful?** Not reported. No confidence intervals or commentary on effect size. **Do conclusions match findings?** Yes.	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Partly. Not clear where treatment sessions take place - could be secondary health setting. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Early help. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Incidence of IPV, quality of parenting **Does the study have a UK perspective?** No. US study.	

21. Thomas R and Zimmer-Gembeck MJ (2012) Parent-child interaction therapy: An evidence-based treatment for child maltreatment. Child Maltreatment 17: 253–66

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: RCT. Part of a larger RCT of Parent-child interaction therapy where participants were allocated to time-variable PCIT (TV/PCIT), standard PCIT (S/PCIT) or waitlist. (Thomas and Zimmer-Gembeck 2011). Data collection: pre- and post- assessment self-report questionnaires, also video-taped pre-assessment. **Description of theoretical approach?** Partly. **How was selection bias minimised?** Randomised. Methods of randomisation not reported. **Was the allocation method followed?** Not reported. **Is blinding an issue in this study?** No blinding. **Did participants reflect target group?** Partly. **Were all participants accounted for at study conclusion?** Yes. See fig. 1. Twenty parents discontinued intervention in S/PCIT group; 27 parents discontinued waiting list group. 'There was no	**Was the exposure to the intervention and comparison as intended?** Not reported. **Was contamination acceptably low?** Not reported. **Did either group receive additional interventions or have services provided in a different manner?** Partly. Waiting list group asked to refrain from family therapy and therapeutic assistance with child behaviour management for the duration of 12 weeks. At the end of 12 weeks, families were offered S/PCIT. ('Families who commenced S/PCIT after the waitlist were not included in the S/PCIT treatment group data of the current study' p257). **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Yes. Validated instruments. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes.	**Does the study's research question match the review question?** Partly. Some of the participants were considered as having engaged in child maltreatment. **Has the study dealt appropriately with any ethical concerns?** No. Not reported. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Homes and clinics. **Does the study relate to at least one of the activities covered by the guideline?** Yes. Targeted activities/interventions to prevent child abuse (early help).	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity score:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
significant difference in attrition rate between S/PCIT and waitlist participants Chi-Square (1, 150) =0.17, p=0 .722.' (p262); 'No differences found between S/PCIT and waitlist participants who completed the RCT and those who did not. No differences between participants who completed or chose not to complete in pre-assessment measures for child behavior, parent stress, depression, child abuse potential, and observational assessment scores. Further, no significant group differences between treatment completers and dropouts were found for child age, parent age, marital status, education level, employment status, or referral source' (p259).	**Were there similar follow-up times in exposure and comparison groups?** Yes. 12 weeks. **Was follow-up time meaningful?** Partly. 12 weeks. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. Similar characteristics between the two groups at baseline (Table 1). **Was intention to treat (ITT) analysis conducted?** Yes. See fig 1. No participants were excluded from analysis. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. No power calculation re sample size. **Were the estimates of effect size given or calculable?** Yes. Estimates of effect size in Cohen's d given. **Were the analytical methods appropriate?** Yes. Analysis of variance (ANOVA) for outcomes; Chi-	**(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **Does the study have a UK perspective?** No. Australian.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	squared tests to calculate differences between intervention and control group. Families who commenced S/PCIT after the waitlist were not included in the S/PCIT treatment group data of the current study. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. **Do conclusions match findings?** Yes.		

22. Thomas R and Zimmer-Gembeck MJ (2011) Accumulating evidence for parent-child interaction therapy in the prevention of child maltreatment. Child Development 82: 177–92

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: RCT. **Description of theoretical approach?** Partly. The authors do not present a clear logic model or theory of change. They report that Parent Child Interaction is an intervention known to reduce child maltreatment and although its effectiveness for parents already involved with child protective services is not clear there are '... many reasons to expect its potential utility' (p177). They go on to discuss the negative impact which	**Was the exposure to the intervention and comparison as intended?** Not reported. **Was contamination acceptably low?** Not reported. **Did either group receive additional interventions or have services provided in a different manner?** Not reported. **Were outcomes relevant?** Yes. Maltreatment outcomes included	**Does the study's research question match the review question?** Yes. The study aimed to examine the effectiveness of a standard Parent Child Interaction Therapy programme for mothers at a high risk of child maltreatment or those with a history of child maltreatment. **Has the study dealt appropriately with any ethical concerns?** Partly. No details on ethical approval or the consent process are given but the authors note in the	**Overall assessment of internal validity:** + The study appears reasonably well designed although significant amounts of methodological details are not provided. In addition, the decision to offer the therapy to the control group after the 12-week assessment point (in itself a very short period) means a significant proportion of the findings are not based on a very robust design. It should also be noted that only 42% of the intervention group

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
harsh and aggressive discipline and communication techniques can have and note that these may escalate to maltreatment. They report that Parent Child Interaction therapy is designed to interrupt these negative cycles by helping parents to understand their behaviour and to manage their interactions with their child. **How was selection bias minimised?** Randomised. No details on the randomisation process are provided. **Was the allocation method followed?** Not reported. Allocation concealment is not reported. **Is blinding an issue in this study?** Blinding not possible. No details on blinding of investigators is given although teacher reports of child behaviour were used as one method of data collection and it seems possible that this group could have been blinded. Due to the nature of the intervention and control (waitlist) it would not have been possible to blind participants and it seems unlikely that blinding of providers would have been possible.	child abuse potential and notification of suspected maltreatment (although data on this measure was not collected for the control group). **Were outcome measures reliable?** Yes. All scales had pre-established reliability and validity. Parenting behaviours were measured via observation but the majority of data was collected via parental reports. **Were all outcome measurements complete?** Partly. Data regarding notification of suspected maltreatment was not collected for the control group. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Partly. At baseline and 12 weeks both groups were assessed; however after 12 weeks the control group were offered therapy due to ethical concerns and no further data from this group was included in the analysis.	limitations section that participation was voluntary. **Were service users involved in the study?** No. No indication that service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. The study evaluates the effectiveness of Parent Child Interaction Therapy on child maltreatment outcomes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Mothers with a history of maltreating their children or those at 'high risk' of doing so. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Partly. The paper does not report settings clearly but it can be assumed that assessments and interventions were conducted in a clinic. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to early help section of guideline. **(For effectiveness questions)**	completed all assessments. **Overall assessment of external validity:** + Only given a + as the study was conducted in Australia. **Overall validity score:** + As the study was conducted in Australia and there are some concerns regarding internal validity (e.g. between group scores could only be analysed at the 12 week point, high rates of attrition in the intervention group, etc.) it is not possible to award a higher overall validity score.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Did participants reflect target group? Yes. Semi-structured interviews confirmed that all participants were at a high risk of child maltreatment. No details on the number of participants who agreed to participate are provided. Participants had been referred by government agencies, including child protection authorities, identified as 'suspects' by a range of professionals, or self-referred. Mothers of children who had suffered from physical maltreatment, emotional maltreatment and neglect were included but those who had suffered from sexual abuse were excluded as Parent Child Interaction Therapy is contraindicated for this group. One mother with substance abuse issues was excluded. **Were all participants accounted for at study conclusion?** Partly. Attrition rates are not reported explicitly. It appears that due to the different methods of data collection drop-out numbers varied for each outcome measure. In addition, the intervention group received a series of assessments which the control group did not. Of the intervention group, only 42%	**Was follow-up time meaningful?** Partly. Between group differences were assessed at baseline and at the 12 week point; after this the control group were offered therapy due to ethical concerns, and data from this group was not included in further analysis. Although this would enable the immediate effects of the program to be detected it is a very short period and would not allow medium to long term effects to be detected. The original control group were then assessed at treatment completion and within group differences were analysed. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Not reported. No details on difference between groups are reported at baseline. **Was intention to treat (ITT) analysis conducted?** Yes. Analysis carried out using last observation carried forward method. This was applied to participants who failed to complete assessment at 12 weeks (study completion) or at completion of programme.	**Are the study outcomes relevant to the guideline?** Yes. Study measures a number of outcomes including child abuse potential and parental stress. **(For views questions) Are the views and experiences reported relevant to the guideline?** Not applicable (not views question). **Does the study have a UK perspective?** No - Australia.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
completed all assessments (a total of three assessments). 71% of families in the control group, completed the single assessment at 12 weeks.	**Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. No power calculation or expected effect sizes are reported. The sample size (n=150) seems small. **Were the estimates of effect size given or calculable?** Not reported. Expected effect sizes are not provided. **Were the analytical methods appropriate?** Yes. ANOVA and chi-square analysis were used. Clinical significance and reliable change indices were calculated when between-group differences were found. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. Effect sizes using Cohen's d are given. Confidence intervals are not provided. **Do conclusions match findings?** Yes. Although when discussing child abuse potential scores the authors seem to overemphasise the reductions detected within the intervention group at treatment		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	completion - there were no significant between group differences at the 12 week assessment point.		

23. Zielinski DS, Eckenrode J, Olds DL (2009) Nurse home visitation and the prevention of child maltreatment: Impact on the timing of official reports. Development and Psychopathology 21: 441–53

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Methodology: RCT. **Description of theoretical approach?** Partly. No description of theoretical approach, although review of existing evidence base is given. **How was selection bias minimised?** Randomised. Abstract reports that families were 'randomly assigned' (p441), however this is not described in the main text. **Was the allocation method followed?** Not reported. **Is blinding an issue in this study?** Blinding not possible. Blinding not possible due to nature of intervention. **Did participants reflect target group?** Partly. Study reports	**Was the exposure to the intervention and comparison as intended?** Not reported. **Was contamination acceptably low?** Not reported. **Did either group receive additional interventions or have services provided in a different manner?** Not reported. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Partly. Study utilises data from Child Protective Services records, focusing on substantiated reports of abuse and neglect. Unclear how reliable these data are, and this is not discussed in the paper. Often these records under-represent true rates of abuse and neglect, and are susceptible to surveillance bias of families already	**Does the study's research question match the review question?** Yes. Research question is on effectiveness of an early help intervention. **Has the study dealt appropriately with any ethical concerns?** Partly. Mentions informed consent, but does not indicate whether ethical approval was sought. **Were service users involved in the study?** No. Service users involved as participants, but not in design or interpretation of results. **Is there a clear focus on the guideline topic?** Yes. Focus is on early help. **Is the study population the same as at least 1 of the**	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Not ++ as not a UK study. **Overall validity score:** + Moderate quality study, with RCT design and substantial follow-up period. However, key limitations included heterogeneity of target population - not all parents had risk factors for abuse and neglect. Also, question reliability of using CPS data alone and the limitations of this are not discussed in the study. Analysis of onset of maltreatment report using two arbitrarily defined time periods is also questionable.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
that '… to avoid creating a program stigmatized as being exclusively for the poor, and to ensure that treatment differences because of risk-level could be examined, any woman who asked to participate and had no previous live birth was accepted in to the study' (p443). This means that not all participants were 'at-risk' - however 85% of the sample had at least 1 of the 3 risk characteristics of young age (<19), single-parent status or low socio-economic status. **Were all participants accounted for at study conclusion?** Yes.	in contact with services. Other studies have also reported difficulty in locating Child Protective Services data due to errors in recording names and dates of birth. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Partly. This study is one of a series examining difference outcomes. This one focuses on the timing of maltreatment. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Yes. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Not reported. **Was intention to treat (ITT) analysis conducted?** Not reported. Unclear whether all 400 original dyads completed treatment. **Was the study sufficiently powered to detect an intervention**	**groups covered by the guideline?** Yes. Study population is caregivers of children at risk of abuse and neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Intervention delivered in participant homes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to early help. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Study measures maltreatment outcomes. **Does the study have a UK perspective?** No. US study.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	effect (if one exists)? Not reported. No power calculation given, and no estimates of effect sizes. **Were the estimates of effect size given or calculable?** Yes. Hazard ratios (exp B) reported. **Were the analytical methods appropriate?** Yes. Use of non-parametric test (Cox regression) means no assumption of underlying distribution. However, query use of p<0.10 significance criterion. Also query analysis by both continuous time AND two time period (age 0–4 and 4–15) - seems an arbitrary categorisation. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes, p values for regression coefficients are given. **Do conclusions match findings?** Yes.		

Review question 9 – Findings tables

1. Barlow J, Simkiss D, Stewart-Brown S (2006) Interventions to prevent or ameliorate child physical abuse and neglect: Findings from a systematic review of reviews. Journal of Children's Services 11: 6–28

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To '… summarise the available evidence from systematic reviews about the effectiveness of interventions to prevent or treat child physical abuse and neglect' (p6). and '… to identify existing systematic reviews of studies of the effectiveness of targeted or indicated interventions for parents that aimed to prevent, reduce or ameliorate incidents of physical abuse or neglect, in order to identify "what works"' (p7). **Methodology:** Systematic review - A systematic review of 15 systematic reviews, published between 1988 - 2005, 10 of which focused on targeted interventions for at-risk families (5 reviews focus on 'indicated' interventions where abuse or neglect has already occurred).	**Participants:** • Children and young people. Parents at risk of abusing and neglecting their children. • Caregivers and families. Parents at risk of abusing or neglecting their children. **Sample characteristics:** • Age - Young children, no details. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief- Not reported. • Disability - 1 review involved parents with intellectual disabilities (e.g. IQs of <80. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - All families 'at-risk', some families of low socioeconomic status. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - No details.	**Effect sizes - Incidence of abuse and neglect:** A - Home visiting (5 reviews) Effect sizes not available for all reviews. Reported effect sizes included: - moderate effect size on unintended childhood injury (ES=0.74) (Roberts et al. 2006, review of 11 RCTs); moderate effect size in relation to reduction of abuse and neglectful acts (ES=0.26) (Geeraert et al. 2004, meta-analysis of 40 studies). Some reviews reported percentage change or other statistics, including: - Relative risk of abuse and neglect ranged from, 0.44 to 11.90 in favour of intervention, 3 results statistically significant (Macmillan et al. 1994, review of 16 RCTS including home visiting and parent training) - 40% change in rates of abuse and neglect in favour of intervention group (Bilukha et al. 2005, review of 22 controlled and uncontrolled studies) - 13% change in rate of out of home placement from baseline (Bilukha 2005) - 4/9 trials showed frequency of occurrence lower in visited group BUT 5/9 trials showed frequency of abuse was higher (Roberts et al. 1996, review of 11 RCTs) Reported statistics for some reviews are unclear, for example: - ES=0.44 (Gray & Halpern, 1988) - unclear what outcome measure this refers to. B - Parenting programmes (3 reviews relating to targeted populations). Relative risk of abuse and neglect ranged from, 0.44 to 11.90 in favour of intervention, 3 results statistically significant (Macmillan et al. 1994, review of 16 RCTS including home visiting and parent training). Gray and Halpern (1988, review of 48 studies) report effect size of 0.42 for 'parenting education classes', however unclear what outcome this refers to.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity score:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Country: Range of countries - mostly US studies. **Source of funding:** Not reported.	• Intervention number - Not reported. • Sample size - Not reported. • Systematic reviews (number of studies) - 15 reviews were considered in total, 10 of which met our PICO (focused on targeted rather than indicated interventions). These broke down as follows: Home visiting - 5 reviews, parenting programmes – 3 reviews, multi-modal interventions – 2 reviews, intensive family preservation services – one review, social support and other interventions – 2 reviews. • The 10 systematic reviews on targeted interventions were based on a total of 509 studies non RCTs and RCTs. **Intervention category:** • Home visiting (five reviews). • Parenting programmes (three reviews). • Multi-modal interventions (two reviews). • Intensive family preservation services (one review). • Social support and other interventions (two reviews). **Intervention:** Five main groups of interventions reviewed: home visiting programmes, multimodal interventions,	C - Multimodal interventions (2 reviews related to targeted interventions) No effect sizes. D - Intensive family preservation services (1 review related to targeted interventions) Dagenais et al. (2004, review of 27 studies). 16 of 27 studies reported placement rates - children who received programme services were placed almost as often as children in control groups (R2=0.008). 17 of 27 studies reported measurements focused on families and children but insufficient data to calculate effect sizes. Available information shows impact on abuse and neglect. E - Social support and other interventions (2 reviews related to targeted interventions) One review found an overall mean effect size for 'at risk' parents of 0.11 on child abuse and neglect. **Effect sizes - Risk of abuse and neglect:** A - Home visiting (5 reviews) Effect sizes for risk reduction reported in Geeraert et al (2004) (review of 40 evaluation studies): - child functioning ES 0.23, - atmosphere ES 0.30, - parent management ES 0.36, - physical ES 0.28, - psychological ES .25, - as parent ES .33, - family function 0.33, - material situation 0.38, - network 0.25. (Geeraert 2004). (Table 2). B - Parenting programmes (3 reviews relating to targeted populations) No data on risk of abuse and neglect. C - Multimodal interventions (2 reviews related to targeted interventions) No effect sizes.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	parenting programmes, intensive family preservation, and social support and other interventions. Home visiting programmes include the Healthy Families America programme, Rooming-in Projects. 'The home visiting programmes evaluated as part of the included reviews were typically provided to high-risk populations of parents (i.e. targeted as opposed to indicated) and comprised a programme of structured visits on a one-to-one basis in the home, delivered by either professionals or specially trained volunteers over an extended period of time. Many of these programmes commenced antenatally or during the immediate postnatal period, and were of variable intensity and duration' (p20). Multimodal interventions were often community based and '… typically comprise family support, preschool education or childcare and community development' (p20). This includes the Project 12 -Ways, '… an eco-behavioural multi-component initiative comprising (a) a behavioural parenting programme, (b) stress reduction, problem-solving and assertiveness training, and (c) single parent services (Thomlison 2003)' (p20). 'A range of early pre-	D - Intensive Family Preservation No effect sizes. E - Social support and other interventions (2 reviews related to targeted interventions) No effect sizes. **Effect sizes - Quality of parenting and parent-child relationships:** A - Home visiting (5 reviews) No effect sizes reported. B - Parenting programmes (3 reviews relating to targeted populations) Feldman et al. (1994, review of 20 mixed methods studies) reports mean improvements in parent outcomes of 63%, and 55% at follow-up. Unclear how this relates to differential improvement between intervention and control. Gray & Halpern, (1998, review of 48 studies) found moderate effect sizes for improvements in knowledge about parenting (ES=1.0), observed parenting behaviour (ES=0.55) but little benefit in terms of attitudes (ES=0.21) or self-report behaviour (ES=0.00). C - Multimodal interventions (2 reviews related to targeted interventions) No effect sizes. D - Intensive family preservation services (1 review related to targeted interventions) Dagenais et al. (2004, review of 27 studies). Seventeen of 27 studies reported measurements focused on families and children but insufficient data to calculate effect sizes. Available information shows impact on family functioning, family support network, family environment, parental disposition.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	ventive interventions, including hospital-based perinatal programmes, support groups, perinatal coaching with home visiting or support group and agency counselling (Gray & Halpern 1988)' (p21). Parenting programmes are typically characterised by: - Brief duration (up to 30 weeks). - Structured interventions delivered either on a 1:1 basis or in groups. - Principal aims to change parenting practice or provide '… new ways of behaving with their child (e.g. basic childcare, safety, nutrition, problem-solving, positive interactions and child behaviour management)...' (p21). Intensive family preservation are usually short-term, home-based interventions which can include '… family focused therapy, support services, behaviour modification, parenting support and life skills training) for families whose children are at risk of out-of-home placement' (p22). The principal aim is to improve '… the safety of the children and family functioning in order to avoid the placement of the children in substitute care' (p22).	E - Social support and other interventions (2 reviews related to targeted interventions) 1 review (Clark 2000, review of 282 studies) found positive effects of programmes on parents with following effect sizes: - Didactic programmes (ES=0.49) - Supportive programmes (ES=0.21) - Combined approach (ES=0.23). Overall mean effect sizes for 'at risk' parents were; 0.14 on parental knowledge and attitudes; 0.26 on dyadic intervention; and 0.27 on caregiving. (Clark 2000). One review (Gray and Halpern 1998, review 48 studies) showed that social support interventions were not effective in changing parenting attitudes (ES =0.21) or behaviour (ES=0.00). **Effect sizes - Children and young people's health and wellbeing outcomes:** A - home visiting (5 reviews) No effect sizes reported. B - Parenting programmes (3 reviews relating to targeted populations) Feldman et al. (1994, review of 20 mixed methods studies) reports mean improvements in child outcomes of 44%, and 39% at follow-up. Unclear how this relates to differential improvement between intervention and control. C - Multimodal interventions (2 reviews related to targeted interventions) No effect sizes. D - Intensive family preservation services (1 review related to targeted interventions) Dagenais et al. (2004, review of 27 studies). Seventeen of 27 studies reported measurements focused on families and children but insufficient data to calculate effect sizes. Available	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Social support and media-based intervention 'The programmes were divided into three categories: (a) "didactic" programmes (i.e. structured interventions with a curriculum, and sequenced activities aimed at shaping or changing participating parents' behaviour, knowledge or outlook); (b) 'supportive' programmes (i.e. loosely structured interventions focused on parent well-being, custom-tailored); and (c) programmes combining both approaches (a and b). Programmes were delivered using both individual and group formats, and for varying periods of duration and frequency (Clark 2000; Gray 1988)' (p22). **Comparison intervention:** Not reported. **Outcomes measured:** • Incidence of abuse and neglect - Abuse and neglect; abuse reduction. 'Documented or reported abuse or neglect, or predictors of abusive parenting such as parenting attitudes and practices, anger and stress levels' (p8). 'Most reviews used predictive measures (i.e. those known to be predictive of abuse and neglect) to evaluate effectiveness, such as	information shows impacts on child performance; delinquency; relationships with peers; and child symptomatology. E - Social support and other interventions (2 reviews related to targeted interventions) 1 review found an overall mean effect size for 'at risk' parents of 0.09 on child development. (Clark 2000). **Effect sizes - Caregiver/parent health and wellbeing outcomes:** A - home visiting (5 reviews) No effect sizes reported B - Parenting programmes. Parent outcomes: Mean per cent improvement across all studies: 63%; Mean per cent improvement at follow-up: 55%; Child outcomes: Mean per cent improvement across all studies: 44%; Mean per cent improvement at follow-up: 39%. (Feldman 1994). C - Multimodal interventions (2 reviews related to targeted interventions) No effect sizes. D - Intensive family preservation services (1 review related to targeted interventions) No effect sizes. E - Social support and other interventions (2 reviews related to targeted interventions) No effect sizes. **Effect sizes - Satisfaction with services:** Not reported. **Effect sizes – Other:** Home visiting is associated with improvements in rates of breastfeeding (Elkan 2000) (Table 2).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	assessments of the home environment, maternal mental health and parenting attitudes and practices' (p20). No details how these are measured. • Risk of abuse and neglect - Risk of out-of-home placement. • Quality of parenting and parent-child relationships - family functioning; family support network; family environment; parental disposition; parent training. 'Most reviews used predictive measures (i.e. those known to be predictive of abuse and neglect) to evaluate effectiveness, such as assessments of the home environment, maternal mental health and parenting attitudes and practices' (p20) No details how these are measured. • Children and young people's health and wellbeing outcomes - Injury/ingestion/trauma, child performance; delinquency; relationships with peers; child symptomatology. No details how these are measured. • Caregiver/parent health and wellbeing outcomes - parental knowledge and attitudes, care giving. No details how these are measured. • Satisfaction with services - not reported. • Service outcomes - rates of out-of-home placement.	**Narrative findings – effectiveness:** Based on the 10 reviews targeted at at-risk families. **A - Home visiting programmes (5 reviews):** Incidence of abuse and neglect: 2 reviews provided weak evidence of effectiveness (1 only if results were controlled for detection bias). The remaining reviews did not measure incidence of abuse and neglect in relation to home visiting, in 2 cases this was due to problem of detection bias. Risk of abuse and neglect: One review of home visiting provided evidence that home visiting was effective in improving '… a range of outcomes associated with abuse and neglect…' (p20) including parenting skills, attitudes and behaviours, home environment and frequency of unintentional injury (Elkan et al. 2000). One review which did not distinguish between home visiting and parenting programmes found evidence for moderate impact for predictive outcomes. Overall, the authors concluded in regards to the effectiveness of home visiting programmes that the '… evidence that they can prevent abuse and neglect remains equivocal because of the issue of surveillance bias' (p20). **B - Parenting programmes (3 reviews relating to targeted populations)** None of the included reviews looked at effectiveness of parenting programmes in relation to incidence of abuse and neglect. Two reviews provided evidence that programmes have a positive impact on parents' knowledge and behaviour (Feldman 1994; Gray & Halpern 1988). One study did not provide data. One	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Follow-up: Not reported.	review showed a moderate overall impact (ES: 0.42), with large improvements in parents' knowledge about parenting (ES:1.0) and observed parenting behaviour (ES: 0.55) but little benefit in terms of attitudes (ES: 0.21) or self-report behaviour (ES:0.00). (Gray 1988, very dated) A parenting programme directed at parents with intellectual disabilities (IQ<80) showed improvements in one or more parenting skills (Feldman 1994). Another review did not provide any data. (MacMillan 1994) (p21). Overall, effectiveness of parenting programmes: '… evidence about their effectiveness is more uniform. This may reflect the greater uniformity in the content of such programmes (many are behavioural or cognitive behavioural) and in their frequency and duration (many are provided over the course of 12 weekly sessions). The included reviews identified some evidence to support the use of parenting programmes based on approaches such as cognitive behavioural therapy, parent–child interaction therapy and other well-recognised models such as the Webster-Stratton Incredible Years series to improve some aspects of parent, child and family functioning, both preventively and therapeutically' (p24). **C - Multimodal interventions (2 reviews related to targeted interventions)** One review assessed Project 12-ways and results reported as part of other reviews in this study (Thomlinson 2003). Another review found that outcomes such as parental knowledge, attitudes and behaviour – both self-reported and observed) showed that while some of these interventions were moderately effective (e.g. hospital-based perinatal programmes – ES:0.34; perinatal coaching with home vis-	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		iting – ES0.29; and agency counselling – ES:0.38) others were ineffective (e.g. perinatal coaching with support group, and support groups alone) (Gray and Halpern 1988, dated) Overall effectiveness of multimodal interventions: uncertain. **D - Intensive Family Preservation Services (IFPS) (1 review**): One review showed no evidence to support the use of IFPS in reducing out-of-home placements but reported significant improvements in family functioning; parental disposition; children's performance; delinquency; relationships with peers; child symptomatology; and maltreatment after the intervention (Dagenais 2004). Overall effectiveness of IFPS – 'Evidence examined by this review were affected by surveillance bias and were not therefore able to demonstrate an effect on out-of-home placement. Results did show reasonable evidence of their effectiveness in improving a range of outcomes associated with abuse and neglect, including parent and family functioning' (p24). **E - Social support and other interventions (2 reviews related to targeted interventions)** One review showed that the 'didactic' programmes aimed at supporting parenting were more effective (ES:0.49) than 'supportive' programmes (ES:0.21) or those offering a combined approach (ES:0.23). (Clark 2000). Another review showed that social support interventions were not effective in changing parenting attitudes (ES:0.003) or behaviour (ES:00). (Gray 1988). Overall effectiveness of social support or media-based intervention: '… media-based interventions (i.e. leaflets), support groups and some forms of perinatal	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		coaching with or without support produced particularly low effect sizes' (p25).	

2. Carta JJ, Lefever JB, Bigelow K et al. (2013) Randomized trial of a cellular phone-enhanced home visitation parenting intervention. Pediatrics 132 (Suppl. 2): S167–73

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To examine '… whether mothers in a parenting intervention, Planned Activities Training (PAT), or cellular-phone enhanced version (CPAT) of the intervention would demonstrate greater use of parenting strategies after treatment and at 6 months post-treatment compared with a wait-list control' (pS167). **Methodology:** RCT. **Country:** Not UK – USA. **Source of funding:** Government. Voluntary/Charity	**Participants:** • Children and young people. 3.5 to 5.5 year-old children. • Caregivers and families. Mothers with at least one of the following risk factors: Age<18 at first child's birth, having less than a high school diploma or equivalent, receiving financial assistance, meeting the income eligibility requirement for Head Start of Special Supplemental Nutrition Program for Women, Infants, and Children. **Sample characteristics:** • Age - Children's mean age = 4.56 years. • Sex - All participants appear to be mothers. • Ethnicity - 46% Hispanic, 33% African American, 17% European American, 4% mixed race or other. • Religion/belief - Not reported. Disability - Not reported.	**Effect sizes - Risk of abuse and neglect:** Significant difference of moderate effect size in parenting stress between CPAT and wait list control (WLC) post-intervention (d=0.27). No significant differences between PAT and WLC, or CPAT and PAT. **Effect sizes - Quality of parenting and parent-child relationships:** Both interventions showed significant improvements in use of PAT compared to WLC immediately post intervention, (d=1.13 for CPAT vs. WLC) and (d=0.81 for PAT versus WLC) and at 6-month follow-up (d=0.56 for CPAT vs. WLC and d=0.44 for PAT vs WLC). There was also a moderate effect size for the difference between CPAT and PAT (d=0.38). Both interventions also showed significant improvement in parenting interaction behaviours (KIPS). At 6-month follow up effect sizes were d=0.46 for CPAT vs. WLC and d=0.34 for PAT versus WLC. **Effect sizes - Children and young people's health and wellbeing outcomes:** Significant differences in rates of positive engagement by children in PAT compared to WLC (d=0.29) and CPAT compared to WLC (d=0.43). No significant differences in maternal ratings	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity score:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Families' average estimated annual income was $18,608. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - Wait list control n=116. • Intervention number - Planned Activities Training (PAT) n=142. Cellular phone and text message enhanced Planned Activities Training (CPAT) n=113. • Sample size n=371. **Intervention category:** Parenting intervention. **Intervention:** Planned Activities Training (PAT) is a manualised component of the SafeCare parent training model. It is a relatively brief intervention comprising five sessions. Aims to prevent challenging behaviour and improve parent-child interactions by focusing on ten strategies: Planning activities in	of children's internalising or externalising behaviours. **Effect sizes - Caregiver/parent health and wellbeing outcomes:** None of the groups were significantly different immediately post-intervention on the depression measure. However, at 6-month follow up CPAT mothers showed significantly lower rates of depression than WLC mothers (d=0.31). **Narrative findings – Effectiveness:** Involvement in either CPAT or PAT produced significant differences compared to WLC on a number of measures including: - Use of PAT strategies - Parenting interaction behaviours (KIPS) - Positive engagement by children. CPAT, but not PAT, was significantly better than WLC on the following measures: - Parenting stress - Parental depression (at 6 months, but not immediately post intervention). No significant differences were observed in parental reports of child internalising and externalising behaviours.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	advance, explaining activities, establishing rules and consequences, giving choices, talking about what you are doing, suing positive interaction skills, ignoring minor misbehaviour, giving feedback and providing rewards or consequences. In the first treatment session, mothers identified specific concerns related to play time and created their own PAT checklist based on the 10 PAT strategies. The strategies were modelled by the family coach with the child during play activities, and mothers were then asked to practice them. The coach would then provide positive and corrective feedback on the use of the strategies. Mothers engaged in practice until they had achieved 80% mastery of the strategies on the PAT checklist. In subsequent sessions, the same strategies were taught in different mother-selected activities. In the final session there was additional practice, a progress review and a plan for the future. Intervention provided by a family coach (staff with a BA degree). Mothers in cell-phone enhanced PAT (CPAT) receive the same intervention, but also received a cell phone and cell phone service throughout intervention. This consists of text messages and phone calls between mothers and		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Family Coaches between sessions. Two text messages were sent per day, with one message prompting use of a particular PAT strategy and one inquiring about their implementation of a particular strategy. Family coaches also called mothers once per week between visits. **Comparison intervention:** Participants in wait list control condition participated in all assessments and observations but did not receive PAT or CPAT. **Outcomes measured:** • Risk of abuse and neglect - Parenting stress index, short form (Abidin 1990) - although note this study does not explicitly mention this as a risk factor for maltreatment, however this measure has been used as a proxy for maltreatment in a number of other studies. • Quality of parenting and parent-child relationships - Parenting assessed by: PAT checklist - assesses mothers' use of PAT strategies Keys to Interactive Parenting Scale (KIPs). • Children and young people's health and wellbeing outcomes - Child behaviour measured using Behavior Assessment Scale for Children-2-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Parent Report Scale and Child Behaviour Rating Scale. • Caregiver/parent health and wellbeing outcomes - Parental depressive mood assessed using Beck Depression Inventory. **Follow-up:** Follow-up was immediately post intervention, and at 6 months.		

3. Dawe S and Harnett P (2007) Reducing potential for child abuse among methadone-maintained parents: Results from a randomized controlled trial. Journal of Substance Abuse Treatment 32: 381–90

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The study aimed to evaluate the impact of Parents Under Pressure programme on outcomes such as family functioning (including child abuse potential) in families in which a parent was engaged in a methadone maintenance programme. This was compared to standard care and a '... second brief intervention control group ...' (p381).	**Participants:** • Children and young people. Children (between the ages of 2 and 8) of parents engaged in a methadone maintenance programme. • Caregivers and families. Parents engaged in a methadone maintenance programme with at least 1 child between the ages of 2 and 8. There were two families in which both parents were currently engaged in substance abuse treatment. In each case the mother was deemed to be the primary caregiver and she provided measurement data. **Sample characteristics:**	**Effect sizes - Risk of abuse and neglect:** Child abuse potential measured using the Child Abuse Potential Inventory (Milner 1986 - self-reported): Standard care (z=2.94, p<.001); 'brief intervention' (z=2.526, p<.001); PUP programme (z=4.591, p<.001). Child Abuse Potential Rigidity scale (Milner 1986 - self-reported): Standard care (z=.221, ns); 'brief intervention' (z=.640, ns); PUP programme (z=2.30, p<.001). Change in risk status on the Child Abuse Potential Inventory: High risk to low risk: PUP programme 8 (36%); 'brief intervention' 4 (17%); standard care 0 (0%). Low risk to high risk: PUP programme 0 (0%);	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity score:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Methodology: RCT. **Country:** Not UK - Australia. **Source of funding:** Not reported. No information on funding source is provided.	• Age - Mean age of primary carer in years (sd): 30.33 (6.34). Ages are not reported by group although the authors report that age did not differ by group F (2,61 = 0.111). Mean age of target child in months (sd): 45.9 (17.2). Not reported by group. • Sex - Data were collected from the parent deemed to be the primary caregiver, the majority of whom were mothers (n=54, 84.4% OR 86%). Number of target children who were male: n=39 (60.9%). • Ethnicity - Not reported. Participants who could not understand or read English were excluded from participation. • Religion/belief - Not reported. • Disability - Not clear - a number of participants were in receipt of a disability pension (n=7, 11%). Not reported by group. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Income source: Paid employment n=15 (23%); unemployment benefits n=3 (5%); disability pension n=7 (11%); sole parenting allowance n=35 (55%); other n=4 (6%). • Type of abuse - Not reported.	'brief intervention' 2 (8%); standard care 8 (42%). Remained high risk: PUP programme 8 (36%); 'brief intervention' 13 (56%); standard care 7 (37%). Remained low risk: PUP programme 6 (27%); 'brief intervention' 4 (17%); standard care 4 (21%). Significant improvement and deterioration using Reliable Change Index: RC + Improved: PUP programme 7 (31%); 'brief intervention' 4 (17%); standard care 0 (0%). RC + Deteriorated: PUP programme 0 (0%); 'brief intervention' 4 (17%); standard care 7 (36%). RC No change: PUP programme 15 (68%); 'brief intervention' 15 (65%) standard care 12 (63%). Parenting stress (self-reported) measured using the Parenting Stress Index Short Form (Abidin, 1990): Standard care (z=0.874, ns); 'brief intervention' (z=0.430, ns); PUP programme (z=2.199, p< 001). **Effect sizes - Children and young people's health and wellbeing outcomes:** Child behaviour (self-reported) measured using the sum of the first four scales of the Strengths and Difficulties Questionnaire (Goodman 1997) to give a total problem score: Standard care (z=1.41, ns); 'brief intervention' (z=0.519, ns); PUP programme (z=2.750, p<.001). Child behaviour – Prosocial behaviour: Standard care (z=1.30, ns); 'brief intervention' (z=1.36, ns); PUP programme (z=2.51, p<.001). **Effect sizes - Caregiver/parent health and wellbeing outcomes:** Parental methadone use was measured using case records: Standard care (z=0.521, ns); 'brief intervention' (z=1.545, ns); PUP programme (z=2.355, p<.001). Parental alcohol use was measured	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Looked after or adopted status - Target child subject to court order: n=7 (10.9%) • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - NB. Two comparison groups. Standard care: Baseline n=19, 6 month assessment n=13. 'Brief intervention': Baseline n=23, six month assessment 20. • Intervention numbers - Baseline: n=22. Six month assessment: n=20. • Sample size – (NB. Two comparison groups). Standard care: Baseline n=19, 6 month assessment n=13. 'Brief intervention': Baseline n=23, 6 month assessment 20. Intervention (PUP): Baseline n=22, six month assessment: n=20. Total sample at baseline n=64. **Intervention category:** Parenting programmes. **Intervention:** Parents Under Pressure: 'The PUP program combines methods for improving parental mood and parenting skills within a multi-systemic framework that takes into account the contextual influences on	using the Alcohol Use Disorders Identification Test (Sanders et al. 1993 - self-reported): Standard care ns; 'brief intervention' ns; PUP programme ns. **Narrative findings – Effectiveness:** Child Abuse Potential Inventory: In the standard care group, child abuse potential *increased* significantly over time. In both the PUP and brief intervention groups, CAP scores decreased significantly over time. Child Abuse Potential Rigidity scale: In the standard care group and the 'brief intervention' groups, scores on the rigidity scale did not change significantly over time. In the PUP group there was a significant reduction in scores on the rigidity scale over time. A higher proportion of those in the PUP programme moved from high to low risk during the course of the intervention, than for the other two conditions. Furthermore a greater proportion of PUP participants showed improvement according to the Reliable Change Index, compared to the other two conditions. However, no statistical testing was conducted on these measures. Child behaviour - Strengths and Difficulties Questionnaire: In the standard care group and the 'brief intervention' groups, total problem scores did not change significantly over time. There was a significant reduction in total problem scores in the PUP group. Child behaviour – Prosocial behavior: No significant differences were found in scores on the prosocial behaviour scale in either the standard care or the 'brief	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	family functioning' (p383). Ten modules delivered over the course of ten to 12 weeks. Sessions last between 1 and 2 hours and are delivered in the home. Treatment begins by assessing families and devising a plan for change. Modules are thematic, may '… continue through treatment …' (p384). and often incorporate mindfulness techniques. Modules include: Module 3 – 'Challenging the notion of an idea parent' to strengthen the parent's view that they are competent. Module 4 – 'How to parent under pressure: increasing mindful awareness' Module 5 – 'Connecting with your child and encouraging good behavior' - teaches skills such as use of praise and reward and play skills Module 6 – 'Mindful child management" teaches non-punitive child management techniques' Module 7 – 'Coping with lapse and relapse" teaches skills to reduce likelihood of lapses to use of alcohol and other drugs' Module 8 – 'Extending social networks' identifying sources of support and extending support networks Module 9 – 'Life skills' - practical advice on diet and nutrition, budgeting, healthy lifestyles and so on Module 10 – 'Relationships' - improving communication between partners.	intervention' group. In the PUP group there was a significant increase in scores on this scale. Parenting Stress Index Short Form: There were no significant changes in scores of parenting stress in either the standard care group or the 'brief intervention' group. In the PUP group there was a significant reduction in scores of parental stress. Parental methadone use: Neither the standard care group nor the 'brief intervention' showed significant changes in methadone dose. Participants in the PUP group showed significant reductions in dosage of methadone over time. Parental alcohol use: There were no significant changes in alcohol use over time in any of the treatment condition groups.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Parental progress is recorded in a workbook. The intervention group may also have received case management (depending on family need) outside of these sessions which involved assistance with accessing services, e.g. accompanying parents on legal or school visits or liaising with social services. Treatment was conducted by 2 clinicians with professional qualifications and experience in treating complex families. **Comparison intervention:** NB. Two comparison groups were used. 'Brief intervention': 2 sessions delivered in the clinic. These were based on traditional parent training skills and were provided by therapists who also delivered the Parents Under Pressure programme. Parents used specially designed workbooks. Standard care: Routine care delivered by clinic staff, i.e. appointments with a prescribing doctor every 3 months and case worker access (e.g. for assistance with benefits, employment and housing). **Outcomes measured:** • Risk of abuse and neglect - Child abuse potential measured using the Child Abuse Potential Inventory (Milner, 1986 - self-reported) and the		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Child Abuse Potential Rigidity scale (a subscale of the former, self-reported). The authors report that this scale is less susceptible to attempts to manipulate scores and that parents with high 'faking good' scores on the main scale have higher rigidity scores. Parenting stress (self-reported) measured using the Parenting Stress Index Short Form (Abidin 1990). • Children and young people's health and wellbeing outcomes - Child behaviour (self-reported) was measured using the Strengths and Difficulties Questionnaire (Goodman 1997). A total problem score was found using the sum of scores on the first four scales (emotional symptoms, hyperactivity, conduct problems, and peer problems). Scores on the fifth scale (prosocial behaviour) are given separately. • Caregiver/parent health and wellbeing outcomes - Parental methadone use was measured using case records. Parental alcohol use was measured using the Alcohol Use Disorders Identification Test (Sanders et al. 1993 - self-reported). **Follow-up:** Assessments conducted at baseline, post-treatment or 3 months and 6 months.		

4. DePanfilis D and Dubowitz H (2005) Family connections: A program for preventing child neglect. Child Maltreatment 10: 108–23

Research aims.	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The study assessed outcomes for families receiving two versions of a programme - 'Family Connections' - specifically designed to prevent child neglect. **Methodology:** RCT with 2 intervention groups: one receiving Family Connections for 3 months (FC3) and one receiving intervention for 9 months (FC9). **Country:** Not UK – USA. **Source of funding:** Government.	**Participants:** • Children and young people. Intervention targeted families that had at least 1 child between the ages of 5 and 11. • Caregivers and families. Intervention primarily delivered to caregivers of children and young people. **Sample characteristics:** • Age - Mean age at baseline 36.9 years, sd 12.2. • Sex - Majority (98.1%) of caregivers were female. • Ethnicity - Majority of caregivers (86.4%) African American. • Religion/belief - Not stated. • Disability - Not stated. • Long term health condition - Not stated. • Sexual orientation - Not stated. • Socioeconomic position - Mean annual income $9,571; mean educational level 10.8 years; 57.8% unemployed. • Type of abuse - Intervention specifically targeted at neglect. • Looked after or adopted status - Not stated. • Unaccompanied asylum seeking,	**Effect sizes - Incidence of abuse and neglect:** No significant between-groups differences in Child Protective Services reports during receipt of intervention (chi-square=0.108, p=0.742) or in the six months following intervention (chi-square=0.115, p=0.177). **Effect sizes - Risk of abuse and neglect:** Risk factors - Caregiver depressive symptoms - Significant main effect of time for full sample (F=18.239, p<0.001). From baseline to case closure FC9 group had fewer depressive symptoms than FC3 (F=3.185, p=0.045) but this difference was not significant 6 months later. Parenting stress - No significant differences between groups on any of the subscales of the parenting stress scale. Difficult Child subscale and Parental Distress subscales- main effect of time for entire sample (F=7.020, p=0.001 and F=7.685, p=0.001 respectively), sustained from baseline, to case closure and at 6-month follow-up. Parent-Child Dysfunctional Interaction subscale - no main group effects or interaction effects. Everyday stress - No significant differences between groups or interaction between group and time. There was a significant reduction in stress in the whole sample across time (F=18.377, p<0.001). Protective factors - Parenting attitudes - 4 subscales of the AAPI were analysed. The Empathy subscale	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity score:** -

Research aims.	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	refugee or trafficked children - Not stated. **Sample size:** • Comparison numbers - FC3 n=62. • Intervention number - FC9 n=63. • Sample size n=125. **Intervention category:** Multi-component intervention. **Intervention:** Family Connections 'works with families in their neighborhoods to help them meet the basic needs of their children, reduce the risk of child neglect, and enhance the overall functioning of the family and children' (p340). Primary theoretical foundation is Bronfenbrenner's (1979) theory of social ecology. The intervention uses a 'home-based, family-centred model of practice' (p340), guided by 9 practice principles: 'community outreach, individualised family assessment, tailored interventions, helping alliance, empowerment approaches, strengths perspective, cultural competence, developmental appropriateness and outcome-driven service plans' (p340). Core components of the intervention are a) emergency assistance, b) home-based family intervention (family	showed a main effect of time (F=3.563, p=0.31) but no main effect of group, or interaction between group and time. Similarly, the AAPI Parent-Child Role Reversal subscale showed a main effect of time (F=16.689, p<0.001) but no main effect of group, or interaction between group and time. There were no significant effects in the Parental Developmental Expectations or Value of Corporal Punishment subscales. Parenting Sense of Competence (PSOC) There was a main effect of time on PSOC (F=9.985, p<0.001) but no significant differences between FC3 and FC9 groups, or interaction between group and time. Family Functioning - Five analyses were conducted on the subscales of the SFI: family health, conflict, cohesion, leadership and expressiveness. In all analyses, there was no effect of time, group or group by time interaction. Social Support - 7 analyses were performed, examining the 6 subscales of the SPS and the total score for the scale. There were significant main effects of time for the Guidance subscale (F=5.924, p=0.004), but no main effect of group, or interaction between group and time. For the Attachment subscale, there was a main effect of group (F=6.682, p=0.011) but no interaction between group and time. For the Opportunity for Nurturance subscale, there was a significant interaction of time and group in favour of the FC3 group (F=3.808, p=0.025). Physical care of children - Significant main effects of time for full sample found for: CWBS household furnishing (F=8.314, p=0.005) CWBS overcrowding (F=4.980, p=0.028) CWBS household sanitation (F=4.406, p=0.038). No main effect of group, or interaction between group and time for any measure of	

Research aims.	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	assessment, outcome driven service plans, individual and family counseling), c) service coordination with referrals targeted towards risk and protective factors, d) multifamily supportive recreational activities, e.g. dinner gatherings. Logic model suggests that the short term outcomes of the intervention are to increase protective factors (parenting attitudes, parenting competence, family functioning and social support) and to decrease risk factors (parental depressive symptoms, life stress and parenting stress). Most services provided by graduate social work interns, supervised by a faculty member - at least weekly individual supervision plus weekly clinical seminars. The FC3 group received an average of 17 hours intervention over 12 weeks, and the FC9 group received an average of 31 hours over 36 weeks. **Comparison intervention:** Comparison intervention same as treatment intervention - but a lower 'dosage' - three months compared to nine months.	physical care. Psychological care of children - Significant main effects of time for full sample for: CWBS mental health care (F=5.961, p=0.016) CWBS parental teaching/stimulation (F=9.008, p=0.004). No main effect of group, or interaction between group and time for any measure of psychological care. **Effect sizes - Children and young people's health and wellbeing outcomes:** Child Behaviour Checklist - internalising Significant main effect of time for the whole sample (F=5.744, p=0.004). There was a significant interaction between group and time in favour of the FC9 group for the CBCL - internalising score (F=3.105, p=0.049). The FC9 group showed greater improvements on this scale than FC3. CBCL - externalising Significant main effect of time for whole sample (F=17.433, p<0.001). No significant main effects of group, or interactions between group and time. **Narrative findings – Effectiveness:** There were no significant between-groups differences in Child Protective Services reports during receipt of intervention or in the 6 months following intervention. Parents in both the FC3 and FC9 intervention groups showed significant improvements over time on a number of risk factors for abuse and neglect including: - Caregiver depressive symptoms - Difficult Child and Parental Distress subscales of the Parenting Stress Index - Everyday stress. However, this was not more pronounced in the higher dosage group. Both groups also showed improvements in protective	

Research aims.	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Outcomes measured: • Incidence of abuse and neglect - Measured via official child abuse and neglect reports, from birth until 6 months after the intervention had ended. Coded based on timing relative to intervention, and substantiation status of reports. • Risk of abuse and neglect - Three risk factors: - Caregiver depressive symptoms - measured via Center for Epidemiologic Studies-Depression Scale (CES-D) (Radloff 1977) - Parenting stress - assessed using Parenting Stress Index Short Form (PSI/SF) (Abidin 1995) - Everyday stress measured using the Every Day Stressors Index (ESI) (Hall et al. 1985). Four protective factors: - Parenting attitudes - measured via Adult-Adolescent Parenting Inventory (AAPI) (Bavolek 1984) - Parenting sense of competence - assessed using Parenting Sense of Competence Scale (PSOC) (Gibaud-Wallston & Wandersman 2001) - Family functioning - assessed using Self-Report Family Intervention (SFI) (Beavers et al. 1985) - Social support - assessing via Social Provisions Scale (SPS) (Russell & Cutrona 1984). Child safety Assessed via observation of physical and psychological care (Child Well Being	factors including: - Empathy and Parent-Child Role Reversal scales of the AAPI Parenting attitudes scales - Parenting sense of competence - Guidance subscale of the social support scale, However, this was not more pronounced in the higher dosage group. Improvements were also seen in both FC3 and FC9 groups in: - CWBS household furnishing - CWBS overcrowding - CWBS household sanitation - CWBS mental health care - CWBS parental teaching/stimulation. However, this was not more pronounced in the higher dosage group. The researchers reports that these findings are in line with 'other studies of home-based interventions with high-risk families' (p119). However, the lack of a 'no services' or 'care as usual' control group makes these findings less robust. However, contrary to the investigators' hypothesis, there were few significant differences between the 3-month and 9-month intervention groups. This included no significant differences in Child Protective Services reports between the groups. The investigators suggest a number of possible explanations for this including: - 3-month intervention may have been delivered with greater intensity - Community referrals made at the end of the intervention may have meant the two conditions were not as different as intended - Differing attrition levels in the 2 conditions.	

Research aims.	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Scales et al. 1986). • Children and young people's health and wellbeing outcomes - Child behaviour measured by caregiver report of externalising and internalising child behaviour problems using the Child Behavior Checklist (Achenback 1991). **Follow-up:** Follow-up at case closure (3 or 9 months), and 6 months post case closure.		

5. DePanfilis D, Dubowitz H, Kunz J (2008) Assessing the cost-effectiveness of Family Connections. Child Abuse and Neglect 32: 335–51

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: 'To assess the cost-effectiveness of two alternate forms of Family Connections (FC), a child neglect prevention program, in relation to changes in risk and protective factors and improvements in child safety and behavioural outcomes' (p335). **Methodology:** RCT with 2 intervention groups: 1 receiving	**Participants:** • Children and young people. Targeted families had at least one child between the ages of 5 and 11. • Caregivers and families. Intervention primarily delivered to caregivers of children and young people. **Sample characteristics:** • Age - Mean age at baseline 36.9 years, sd 12.2. • Sex - Majority (98%) of caregivers were female. • Ethnicity - Majority of caregivers (86%) African American. • Religion/belief - Not stated.	**Effect sizes – Other:** See DePanfilis et al. (2005) for effect sizes on outcome measures. Cost effectiveness ratio for FC3 group with regard to child behaviour was $337. Cost-effectiveness ratio for FC9 group with regard to child behaviour was $276. **Narrative findings – Effectiveness:** For narrative of effectiveness on outcome measures see DePanfiis et al. (2005). This study found that, for the majority of outcome measures, the FC3 group was more cost effective than FC9. This is because there was no significant difference between the groups in risk and protective factors, or child safety. Both groups showed a similar magnitude of improvement at 6 months follow-up, suggesting that the briefer, and less	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity score:** -

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Family Connections for 3 months (FC3) and 1 receiving intervention for 9 months (FC9). **Country:** Not UK – USA. **Source of funding:** Government.	• Disability - Not stated. • Long term health condition - Not stated. • Sexual orientation - Not stated. • Socioeconomic position - Mean annual income $9,571; mean educational level 10.8 years; 57.8% unemployed. • Type of abuse - Intervention specifically targeted at neglect. • Looked after or adopted status - Not stated. • Unaccompanied asylum seeking, refugee or trafficked children - Not stated. **Sample size:** • Comparison numbers - FC3 n=62. • Intervention number - FC9 n=63. • Sample size n=125. **Intervention category:** Home visiting. **Intervention:** Family Connections 'works with families in their neighborhoods to help them meet the basic needs of their children, reduce the risk of child neglect, and enhance the overall functioning of the family and children' (p340). Primary theoretical foundation is Bronfenbrenner's (1979) theory of social ecology. The intervention uses a 'home-based, family-centred	expensive, intervention was the more cost effective. However, in relation to child behaviour outcomes the FC9 intervention was calculated to be the more cost effective - costing $276 per unit change in child behaviour compared to $337 for FC3.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	model of practice' (p340), guided by nine practice principles: 'community outreach, individualised family assessment, tailored interventions, helping alliance, empowerment approaches, strengths perspective, cultural competence, developmental appropriateness and outcome-driven service plans' (p340). Core components of the intervention are a) emergency assistance, b) home-based family intervention (family assessment, outcome driven service plans, individual and family counseling), c) service coordination with referrals targeted towards risk and protective factors, d) multifamily supportive recreational activities, e.g. dinner gatherings. Logic model suggests that the short term outcomes of the intervention are to increase protective factors (parenting attitudes, parenting competence, family functioning and social support) and to decrease risk factors (parental depressive symptoms, life stress and parenting stress). Most services provided by graduate social work interns, supervised by a faculty member - at least weekly individual supervision plus weekly clinical		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	seminars. The FC3 group received an average of 17 hours intervention over 12 weeks, and the FC9 group received an average of 31 hours over 36 weeks. **Comparison intervention:** Comparison intervention same as treatment intervention - but a lower 'dosage' - three months compared to nine months. **Outcomes measured:** • Incidence of abuse and neglect - Measured via official child abuse and neglect reports, from birth until 6 months after the intervention had ended. Coded based on timing relative to intervention, and substantiation status of reports. • Risk of abuse and neglect - Three risk factors: - Caregiver depressive symptoms - measured via Center for Epidemiologic Studies-Depression Scale (CES-D) (Radloff 1977) - Parenting stress - assessed using Parenting Stress Index Short Form (PSI/SF) (Abidin 1995) - Everyday stress measured using the Every Day Stressors Index (ESI) (Hall et al. 1985). Four protective factors: - Parenting attitudes - measured via		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Adult-Adolescent Parenting Inventory (AAPI) (Bavolek 1984) - Parenting sense of competence - assessed using Parenting Sense of Competence Scale (PSOC) (Gibaud-Wallston & Wandersman 2001) - Family functioning - assessed using Self-Report Family Intervention (SFI) (Beavers et al. 1985) - Social support - assessing via Social Provisions Scale (SPS) (Russell & Cutrona 1984). Child safety Assessed via observation of physical and psychological care (Child Well Being Scales, Magura & Moses 1986). • Children and young people's health and wellbeing outcomes - Child behaviour measured by caregiver report of externalising and internalising child behaviour problems using the Child Behavior Checklist (Achenbach 1991). **Follow-up:** Follow-up at case closure (3 or 9 months), 6 six months post case closure.		

6. Dishion T, Mun Chung J, Drake Emily C et al. (2015) A transactional approach to preventing early childhood neglect: The Family Check-Up as a public health strategy. Development and psychopathology 27: 1647–60

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To investigate whether a home-based visitation intervention, the Family Check-Up (FCU) reduces the risk of child of maltreatment by improving the parent-child relationship in low-income US families. **Methodology:** RCT inc cluster. Families were recruited via referrals from Women, Infants and Children nutritional assistance programme in 3 American areas. Eligible families had children aged 2 years and 2 years 11 months and reported higher than average scores in at least 2 out of 3 domains of risk factor: familial (maternal depression; daily parenting challenges, substance use problems, teen parent status); child-related (conduct problems, high-conflict relationships with	**Participants** Children and young people - Children involved in assessments - aged 2 at baseline. Caregivers and families - Families were invited to participate in our study on the basis of the following inclusion criteria: (a) they had a child between ages 2 years 0 months and 2 years 11 months, and (b) they reported family, socioeconomic, and/or child risk factors for child's future behaviour problems; more specifically, families had to score at least 1 SD above the normative mean in two of the three domains of risk: familial (i.e., maternal depression, daily parenting challenges, substance use problems, or teen parent status); child (i.e., conduct problems and high-conflict relationships with adults); and socio-demographic (i.e., low education achievement and income relevant to WIC criterion). **Sample characteristics** • Age - Age of child at baseline - mean =29.9 months (SD= 3.2). • Sex - Children 49% female (n=180 control, n=182 FCU). No detail given on gender of caregivers in the study, but participation in WIC suggests female.	**Effect sizes** Incidence of abuse and neglect Impact of FCU on neglect variables. A path model for the relationship between intervention group, dyadic positive engagement and neglect measures was constructed. The analysis found no direct effect of FCU upon neglect variables at age 4 follow up. However, the path analysis found that the relationship between FCU and neglect was mediated via impact on dyadic positive engagement, and moderated by level of family adversity at age 2. Using a mean family adversity score, the study found significant indirect effects of FCU on neglect mediated by improved dyadic positive engagement for: affection neglect (point estimate -0.077 95% CI -0.137 to -0.017); monitoring neglect (point estimate -0.034 (95% CI -0.068 to -0.001) and caregiving (point estimate -0.019, 95% CI -0.035 to -0.003). Family adversity at age 2 was found to be significantly associated, at age 4, with affection neglect (B=0.32, SE=0.04, p<.05) caregiving neglect (B=0.13, SE=0.04, p <0.01) and monitoring neglect,(B=0.32, SE = 0.10, p< 0.01) but not with DPE at age 3 (B=-0.08, SE=0.01, p=0.14). Family adversity also significantly modified the association between DPE and affection neglect (B=-0.71, SE= 0.33, p<0.05) and monitoring effect (B=-0.52, SE=0.27, p=0.05). A slope analysis of this moderating effect found that as family adversity index increased, the negative association between DPE at age 3 and affection or monitoring neglect at age 4 became stronger, meaning that families with high and mean adversity scores saw the biggest decrease in affection or monitoring neglect as their DPE score increased, compared to families with low	**Overall assessment of internal validity** + Lack of UK focus **Overall assessment of external validity** + **Overall validity score** + Rationale for data analysis not always clear.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
adults); and socioeconomic (low education achievement and income assistance relevant to WIC criteria). Child maltreatment was not a criteria of eligibility. Of 1666 families approached, 879 were eligible and 731 agreed to take part. Following a home assessment (when child aged 2) participating families were randomly assigned to the intervention condition or control condition, and then assessed at home 3 times further, at ages 3,4 and 5 with assessors blind to the group status of the family. For those in the FCU intervention group, these assessments occurred prior to the 3 annual intervention sessions. Assessments involved videotaped observation of tasks which the participating primary caregiver (PC)-child dyad	• Ethnicity - 46.6% (n=341, control n=170, FCU n=171) European American, 27.6% (n=202; c n=97, FCU n=105) African American, 13.4% (n=98, c n=48, FCU n=50) Hispanic, 9.8% (n=72, c and FCU n=36) biracial and 2.4% (n=18, c n=13, FCU n=5) other (First people/native Hawaiians) - control and interventions similar to each other. • Religion/belief - Not reported • Disability - Not reported • Long term health condition - Not reported • Sexual orientation - Not reported • Socioeconomic position - Primary caregivers education 23.6% (n= 172, c n=92, FCU=80) had less than high school diploma, 41% (n= 310, c n=137, FCU n=163) had a high school education or general education diploma, 35.4% (n=259, c n=135, FCU n=124 had post-high school training. PC income less than $10,000 - 28.7% (n=210, c n=108, FCU n=102; Less than $20,000 - 37.6% (n= 275, c n=139, FCU n=136) $20,000 and over - 32.5% (n=238, c n=113, FCU n=125). Mean number of family members per household was 4.5 (SD 1.63). • Type of abuse – Neglect. • Looked after or adopted status – Not applicable - biological children living with their primary caregivers	adversity scores, where there was no significant association. Quality of parenting and parent-child relationships DPE was measured at age 2 and 3: FCU participants scored significantly higher DPE scores at age 3 than those in the control group, while controlling for baseline DPE (b=0.03, SE=0.01, p<0.01). At age 3 higher DPE scores significantly predicted less affection neglect (B=-2.29, SE=0.15, p<0.001) caregiver neglect (B=-0.56, SE=0.15, p<0.001) and monitoring neglect (B=-1.02, SE=0.36, p<0.01). The effect was seen controlling for same baseline measures at age 2, family adversity at age 2 and FCU direct effect.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
were asked to complete - a free-play task for the child (15 mins); then for both PC and child a clean-up task (5 mins); a delay of gratification task (5 mins) and 4 teaching tasks (3 mins each). This was followed by a second 4 min free-play task for the child and a second 4 min clean up task for both. The assessment concluded with a 2 min task involving 'inhibition-inducing toys' (p1651) and a 20 min meal prep/lunch task. Videos were then later coded by researchers for positive engagement, using the Relationship Affect Coding system, which codes physical and verbal behaviour and affect of both the PC and the child as either positive, negative or neutral, based on facial expressions, tone of voice and nonverbal cues. Software was	• Unaccompanied asylum seeking, refugee or trafficked children - Not applicable. **Sample size** Comparison numbers - n=364 at baseline, n=309 at age 4 follow-up. Intervention numbers - n=367 at baseline, n=315 at age 4 follow-up. Sample size - n=731 at baseline, n=624 at age 4 follow-up. **Intervention** Intervention category – Home visiting. Describe intervention - FCU offers 3 home-based annual sessions, tailored to families following assessment results, comprising of an observation session, an interview session and a feedback session. Families are invited to participate in assessment of observation, interview and feedback to avoid observation assessment bias. Delivered by No detail given other than 'staff' except for interview sessions which are delivered by trained consultants (p1651).		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
used to calculate summary dyadic positive engagement (DPE) scores. While videos from ages, 2, 3, 4, and 5 were coded, only data from ages 2 and 3 was included in the analysis. Child neglect at age 4 was coded using the Coder Impressions Inventory, rating PC-child relationship quality; negative PC-child interaction and family problem-solving skills on a 9 point scale of 1=not at all, 5=somewhat, 9=very much, leading to calculation of scores for 2 variables, affection neglect and monitoring neglect. Caregiving neglect, the home environment, was assessed at age 4 using the Home Observation for Measure of the Environment during a 3 hr visit by an examiner, with the score for 'basic hygiene', measured on a 4-point Likert scale,	Delivered to Primary caregivers and their child (aged 2 at baseline), but with wider family involved at all stages where available, particularly to provide information during observation session. Duration, frequency, intensity, etc. One home-based session every year, for three years. FCU sessions were delivered after annual assessments had taken place. Key components and objectives of intervention An observation session involving wider family is designed to gather a holistic ecological picture of the PC and child relationship. During the interview session, PCs are given the opportunity to raise concerns, particularly those about family issues which impact on the child's wellbeing. The feedback session discusses the observed parenting, the strengths seen and possible areas of change in the family and parenting practices. All three sessions are designed to improve relationships between the primary caregiver and the child, and also with the wider family, thus reducing likelihood of child maltreatment. A motivational interviewing style is used to support primary caregivers to seek		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
(1=not at all true, 2 = hardly true, 3= somewhat true, 4=very true) used as the outcome measure for this study. Levels of family adversity were calculated combining the scores for 8 indicators: 1. PC income below national poverty line; 2. low PC educational attainment (below high school); 3. single parenthood; 4.household overcrowding (4+ children in home/more family members than rooms). 5 household member with criminal conviction; 6. PC with substance abuse problems; 7.Neighbourhood danger status 8.PC with depression. This score, from 0–8 was generated when the child was 2. Multivariate analysis was conducted for the prediction of DPE (as the mediating factor) and other outcome variables, with percentile bootstrapping used to	to improve their identified weak areas. Information about available and evidence-based parenting support is also provided, culminating in an assessment of whether follow up services, of an Everyday Parenting (Dishion et al. 2011) curriculum is required. Location/place of delivery Own homes Describe comparison intervention Received Women, Infants and Children's services as usual. **Outcomes measured** Incidence of abuse and neglect. Child neglect comprising: 1. Affection neglect - measured using Coder Impressions Inventory (Dishion et al. 2004) at ages 2 and 4 2. Monitoring neglect - measured using Code Impressions Inventory (Dishion et al. 2004) at ages 2 and 4 3. Caregiving neglect - measured using Home Observation for Measurement of Environment (Caldwell and Bradley 2003). Quality of parenting and parent-child relationships. Observed dyadic positive engagement - measured using the Relationship Affect Coding System (Peterson et al. 2008) at ages 2 and 3.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
assess mediation effects. **Country:** Not UK. USA. **Source of funding:** Government. National Institute on Drug Abuse.	**Follow-up** Assessments were conducted at baseline, age 2, age 3, age 4 and age 5). However, only data from ages 2 and 3 were used in relation to dyadic positive engagement, and only data from ages 2 and 4 were used in relation to neglect. **Costs** No.		

7. DuMont K, Kirkland K, Mitchell-Herzfeld S et al. (2011) Randomized Trial of Healthy Families New York (HFNY): Does Home Visiting Prevent Child Maltreatment? New York: New York State Office of Children and Family Services

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: Study aims to answer four questions: '1) To what extent is the home visiting process of HFNY consistent with the HFA model? 2) Does home visiting effectively prevent or reduce maltreatment? 3) Does home visiting limit the emergence of precursors of delinquence? 4) Do the long-term benefits of an	**Participants:** • Caregivers and families. **Sample characteristics:** • Age - Baseline, 31.0% of mothers >19 Follow-up mother sample: 31.7% of mothers >19 Follow-up child sample: 32.1% of mothers >19. • Sex - All parent participants appear to be female. Proportion of target children who were female: Baseline = 46.1%; Follow-up mother sample: 46.5%; Follow-up child sample 47.3%.	**Effect sizes - Incidence of abuse and neglect:** **A - Administrative data** 1. Overall sample 1.1 Child welfare reports In the overall sample there were no significant difference in either cumulative rates or cumulative numbers of confirmed reports of abuse and neglect between the intervention and control groups: Mother or target child confirmed subject or victim of CPS report (cumulative rate) - Control 27.10%; Intervention 29.55; Adjusted odds ratio=1.13, non-significant Mother or target child confirmed subject or victim of CPS report (mean number of reports) - Control 0.55; Intervention 0.54; Effect size=-0.01, non-significant. 1.2 Initiation of child welfare services No significant difference in initiation of child welfare services. Control 18.61%; Intervention	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** + **Overall validity score:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
HFA-based home visiting program outweigh its costs?' (p4-5). **Methodology:** RCT. **Country:** Not UK – USA. **Source of funding:** Government.	• Ethnicity - Mother's ethnicity Baseline: White 34.4%; African-American 45.4%; Latina 18.0% Follow-up mother sample: White 34.6%; African-American 47.9%; Latina 15.6% Follow-up child sample: White 34.1%; African-American 49.0%; Latina 15.1%. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Receiving cash assistance: Baseline 36.5% Follow-up mother sample: 37.8% Follow-up child sample: 36.3% • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - Baseline comparison group n=594 Follow-up (Year 7) comparison group n=463. • Intervention numbers - Baseline intervention group n=579 Follow-up intervention group n=463. • Sample size - Baseline comparison group n=594 Follow-up (Year 7)	16.21%; Adjusted odds ratio=0.87, non-significant. 1.3 Foster care placement No significant difference in rates of foster care placement. Control 4.90%; Intervention 4.83%; Adjusted odds ratio=0.87, non-significant. Subgroup analyses were conducted for two subgroups: a 'Reduced Risk of Offending' (RRO) group (n=104) and a High Prevention Opportunity (HPO) group (n=179). 2. RRO subgroup 2.1 Child welfare reports Marginally significant difference in favour of intervention group regarding cumulative rates of confirmed reports of all types of abuse and neglect: Control 60.36%; Intervention 41.51%; Adjusted odds ratio 0.47, p<0.10. Marginally significant difference in favour of intervention group regarding reports where mother was the confirmed subject: Control 57.41%; Intervention 38.18%; Adjusted odds ratio 0.46, p<0.10. Marginally significant difference in favour of intervention group regarding cumulative rates of confirmed reports of physical abuse: Control 13.44%; Intervention 3.25%; Adjusted odds ratio 0.22, p<0.10. Marginally significant difference in favour of intervention group regarding mean numbers of confirmed reports of all types of abuse and neglect: Control 1.63; Intervention 0.96; Effect size -0.35, p<0.10. 2.2 Initiation of child welfare services Significant difference in favour of intervention group regarding initiation of child welfare services: Control 60.02%; Intervention 38.02%; Adjusted odds ratio 0.41, p<0.05. 2.3 Foster placements No significant difference in rates of foster care placement. Control 23.62%; Intervention 17.19%; Adjusted odds ratio 0.67, non-significant. 3. High Prevention Opportunity subgroup 3.1 Child	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	comparison group n=463 Baseline intervention group n=579 Follow-up intervention group n=463. **Intervention category:** Home visiting. **Intervention:** Healthy Families New York aims to provide 'voluntary, comprehensive and intensive home visiting services to expectant or new parents who are identified as being at risk of abusing or maltreating their children' (p6). The service comprises: - bi-weekly visits during the prenatal period - weekly visits until the child is at least 6 months old - periodic visits based on the needs of the family until the child begins school or Head Start. Home visitors provide '… support, education, information and activities designed to promote healthy parenting behaviors and child growth, including proper nutrition, age appropriate behaviors and positive discipline strategies' (p7). They also help mother's access health care and other services. The service is delivered by trained paraprofessionals. **Comparison intervention:** Control group provided with information on, and referral to, appropriate services other than home visiting. **Outcomes measured:**	welfare reports No significant difference in either cumulative rates or cumulative numbers of confirmed reports of abuse and neglect between the intervention and control groups: Mother or target child confirmed subject or victim of CPS report (cumulative rate) - Control 25.03%; Intervention 21.92%; Adjusted odds ratio =0.4, non-significant Mother or target child confirmed subject or victim of CPS report (mean number of reports) - Control 0.49; Intervention 0.31; Effect size= -019, non-significant. 3.2 Initiation of child welfare services No significant difference in initiation of child welfare services. Control 12.3%; Intervention 8.67%; Adjusted odds ratio=0.68, non-significant. 2.3 Foster placements No significant difference in rates of foster care placement. Control 1.76%; Intervention 2.88%; Unadjusted odds ratio 1.47 (incidence too low to support multivariate model), non-significant. **B - Self-report data - mothers** 1. Overall sample Significant differences in favour of intervention group in relation to frequency of non-violent discipline and frequency of serious physical abuse. Non-violent discipline - Control: 45.27; Intervention 49.27, Effect size 0.14. p<0.05. Serious physical abuse - Control: 0.15; Intervention 0.03, Effect size -0.20. p<0.01. Also significant difference in rates of non-violent discipline - Control: 98.6%; Intervention 100.00, Effect size 0.14. p<0.05 No significant differences observed on overall rates (whether it occurs or not) of: Psychological aggression Control: 86.49%; Intervention 87.29%, Adjusted odds ratio 1.18, non-significant. Minor physical aggression Control: 59.17%; Intervention 64.12%, Adjusted odds ratio 1.25, non-	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Incidence of abuse and neglect - Measured using three sources: 1. Administrative records - reports of child maltreatment looking at a) cumulative rates (whether any confirmed reports of abuse or neglect over the time period) and b) cumulative numbers of reports, initiation of family services to avoid a placement and foster care placement 2. Self-reported by mother - using revised Parent-Child Conflict Tactics scale (Straus et al. 1998) using subscales for non-violent discipline, psychological aggression, minor physical aggression, serious physical abuse and neglect. 3. Self-reported by child - using Conflict Tactics Scale-Picture Card Version (Mebert and Straus 2002), looking at non-violent discipline, psychological aggression and minor physical aggression. • Children and young people's health and wellbeing outcomes - The study also measured a number of 'precursors to delinquency' to assess if the intervention had reduced risk of juvenile delinquency. These were measured using: 1. Interview with mothers regarding child's experiences 2. Child Behavior Checklist for Ages 6-18 (Achenbach & Rescorla 2001) 3. Child interview - receptive language skills measured	significant. Serious physical abuse Control: 3.18%; Intervention 1.76%, Adjusted odds ratio 0.55, non-significant. Neglect Control: 16.74%; Intervention 15.77%, Adjusted odds ratio 0.93, non-significant. Also no significant differences observed in frequency of: Psychological aggression Control: 15.21; Intervention: 15.33; Effect size 0.01, non-significant Minor physical aggression Control: 4.51; Intervention: 4.36; Effect size -0.02, non-significant Neglect Control: 0.64; Intervention: 0.53; Effect size 0.05, non-significant 2. HPO subgroup There was a marginally significant difference in favour of the intervention group in overall rate of psychological aggression: Control 91.19%; Intervention 79.74%; Adjusted odds ratio 0.38, $p<0.10$ There was a marginally significant difference in favour of the intervention group in frequency of serious physical abuse: Control 5.47; Intervention 3.10; Effect size -0.34, $p<0.10$. No significant differences observed on overall rates (whether it occurs or not) of: Non-violent discipline - Control: 100%; Intervention 100%, Adjusted odds ratio non-significant. Minor physical aggression - Control: 65.58%; Intervention 64.79%, Adjusted odds ratio 0.92, non-significant. Serious physical abuse - Control: 3.40%; Intervention 3.20%, Adjusted odds ratio 0.55, non-significant. Neglect - Control: 12.53%; Intervention 17.07%, Adjusted odds ratio 1.39, non-significant. Also no significant differences observed in frequency of: Non-violent discipline - Control: 45.14; Intervention: 43.30; Effect size -0.06, non-significant. Psychological aggression - Control: 12.99; Intervention: 9.93; Effect size -0.23, non-significant Neglect Control: 0.28; Intervention: 0.27; Effect size 0.01, non-significant. **C - Self-report data - Children**	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	using Peabody Picture Vocabulary Teset (Dunn & Dunn 2007) 4. Loneliness and Social Dissatisfaction Questionnaire (Cassidy & Asher 1992) 5. Anti-social tendencies assessed using questions adapted from Seattle Social Development Project (Hawkins 2003) 6. Automated Delay of Gratification task, which show associations with social and academic outcomes (Mischel et al. 1989). **Follow-up:** This study reports the 7-year follow-up. Previous waves of the study also followed up at Years 2 and 3.	1. Overall sample There was a significant difference in favour of the intervention regarding prevalence of minor physical aggression: Control 77.23%; Intervention 70.79%; Adjusted odds ratio 0.74, p<0.05. There was no significant difference in: Prevalence of non-violent discipline - Control 96.90%; Intervention 97.80; Adjusted odds ratio 1.33, non-significant. Prevalence of psychological aggression - Control 85.14%; Intervention 84.47; Adjusted odds ratio 1.00, non-significant. There were no significant differences in: Frequency of non-violent discipline - Control 4.02; Intervention 4.03; Effect size 0.01; non-significant. Frequency of psychological aggression - Control 2.68; Intervention 2.78; Effect size 0.05; non-significant. Frequency of minor physical aggression: Control 2.35; Intervention 2.27; Effect size -0.04; non-significant. 2. HPO subgroup There was no significant difference in: Prevalence of non-violent discipline - Control 96.70%; Intervention 95.20; Adjusted odds ratio non-significant. Prevalence of psychological aggression - Control 87.13%; Intervention 84.93%; Adjusted odds ratio 0.87, non-significant. Prevalence of minor physical aggression: Control 81.57%; Intervention 75.70%; Adjusted odds ratio 0.79, non-significant. There was no significant difference in: Frequency of non-violent discipline - Control 3.93; Intervention 3.60; Effect size -0.20, non-significant. Frequency of psychological aggression - Control 2.92; Intervention 3.00; Effect size 0.03; non-significant. Frequency of minor physical aggression: Control 2.59; Intervention 2.63; Effect size 0.02; non-significant. **Children and young people's health and wellbeing outcomes –**	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		1. Interview with mothers regarding child's experiences: No effect sizes. 2. Child Behavior Checklist for Ages 6-18 (Achenbach & Rescorla, 2001) Whole sample: Rule breaking behaviours=ns; Aggressive behaviours = ns; Anxious depressed=ns; Withdrawn depressed = ns; Social problems=ns. HPO subgroup: Rule breaking behaviours=ns; Aggressive behaviours=ns; Anxious depressed=ns; Withdrawn depressed=ns; Social problems=ns. 3. Child interview - receptive language skills measured using Peabody Picture Vocabulary Teset (Dunn & Dunn 2007). Whole sample Receptive vocabulary=ns HPO subgroup Receptive vocabulary=significant difference in proportion with below average receptive vocabulary, in favour of intervention, ES=0.43. 4. Loneliness and Social Dissatisfaction Questionnaire (Cassidy & Asher 1992): Whole group and HPO subgroup - social isolation=ns; Whole group and HPO subgroup - 'ever bullied by others'=ns. 5. Anti-social tendencies assessed using questions adapted from Seattle Social Development Project (Hawkins 2003). Whole group and HPO subgroup - bullying activities=ns; Whole group and HPO subgroup - deviant activities=ns. 6. Automated Delay of Gratification task, which show associations with social and academic outcomes (Mischel et al. 1989). Whole group and HPO subgroup - all self-regulation measures=ns. 7. Risk for poor school outcomes: Whole sample: % participating in a gifted programme - $p<0.01$; % receiving remedial services=ns; % receiving special education=$p<0.10$; % repeating a grade=ns; % skipping	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		school more than once = ns; % skip school often=p<0.01. HPO subgroup: % participating in a gifted programme - p<0.10; % receiving remedial services=ns; % receiving special education=ns; % repeating a grade=p<0.10; % skipping school more than once=ns; % skip school often=ns. **Narrative findings – Effectiveness:** 1. Administrative data on abuse and neglect 1.1 Reports of abuse and neglect. In the overall sample there were no significant difference in either cumulative rates or cumulative numbers of confirmed reports of abuse and neglect between the intervention and control groups. Additional analyses conducted within the study suggested that this may be due to detection bias in the intervention group: because they were already in contact with services, this group was more likely to be reported to CPS. In the RRO subgroup, marginally significant differences in favour of the intervention group were observed regarding cumulative rates of confirmed reports of all types of abuse and neglect; of reports where mother was confirmed subject, and reports of physical abuse. In the HPO subgroup there were no significant differences in either cumulative rates or cumulative numbers of confirmed reports of abuse and neglect between the intervention and control groups. 1.2 Initiation of child welfare services. No significant differences found in the overall sample or HPO subgroup. A Significant difference in favour of intervention group regarding initiation of child welfare services was found in the RRO subgroup.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		1.3 No significant differences in rates of foster placement were found in intervention compared to control in overall sample, RRO subgroup or HPO subgroup. 2. Self-report data parenting behaviours - mothers Significant differences were found in the overall sample in favour of the intervention group in relation to frequency of non-violent discipline and serious physical abuse. No significant differences observed in overall rate of serious physical abuse. No significant differences were observed on overall rates or frequency of psychological aggression, minor physical aggression or neglect. In the HPO subgroup there was a marginally significant difference in favour of the intervention group in overall rate of psychological aggression, but not its frequency. A marginally significant difference was found in favour of the intervention for frequency of serious physical abuse, but not its overall prevalence. No significant differences were observed on overall rates or frequency of non-violent discipline or neglect. 3. Self-report data parenting behaviours – children In the overall sample, there was a significant difference in favour of the intervention regarding prevalence of minor physical aggression: There was no significant difference in prevalence of non-violent discipline or psychological aggression. There was no significant difference in frequency of non-violent discipline, psychological aggression or minor physical aggression. In the HPO subgroup there were no significant differences in any of the above measures. 4. Measures of risk factors for delinquency - In the whole sample, there were few significant differences between the treatment and control groups in any of the measures of risk factors for delinquency, which in-	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		cluded the Child Behaviour Checklist, measures of receptive vocabulary, social isolation and deviant behaviours. In the whole sample, a significant difference was observed in favour of the intervention in terms of participation in a 'gifted' programme at school ($p<0.01$), and proportion skipping school often ($p<0.01$). A marginally significant difference was also observed in terms of proportion in receipt of special education ($p<0.10$). For the HPO subgroup, there was a significant difference in the proportion of young people in favour of the treatment condition in terms of below average receptive vocabulary score ($p<0.05$) and marginally significant differences in participation in a gifted programme ($p<0.10$) and proportion of young people repeating a grade ($p<0.10$).	

8. Green BL, Tarte JM, Harrison PM et al. (2014) Results from a randomized trial of the Healthy Families Oregon accredited statewide program: Early program impacts on parenting. Children and Youth Services Review 44: 288–98

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: '(1) What short-term program effects can be detected at children's 1-year birthday? In particular, compared to control families: (a) Do parents in the Healthy Families Oregon (HFO) group report more positive parenting behaviors and	**Participants:** • Caregivers and families. First time parents with an infant under 3 months scoring positively on any 2 risk factors of the New Baby Questionnaire. This is adapted from the Hawaii Health Risk Indicators Instrument (Duggan 2004). This measure assesses family risk using a range of criteria: mothers under the age of 19, delayed prenatal care, poor en-	**Effect sizes - Risk of abuse and neglect:** Marginally significant effect on parenting stress in HFO compared to control mothers ($F(1,759)=3.621$, $p=0.057$). **Effect sizes - Quality of parenting and parent-child relationships:** Note: Some reported statistics do not appear in tables as advised in text. Attitudes regarding corporal punishment: No significant difference. However, for highest risk families, HFO families significantly less likely to endorse corporal punishment (effect sizes unclear). Parent–child positive activities: Ap-	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity score:** -

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
skills compared to families in the control group? (b) Do parents in the HFO group report lower parenting stress, less depressive symptomatology, and more positive family functioning compared to families in the control group? and (c) Do children in the HFO treatment group experience more supports for healthy development, specifically increased breastfeeding and increased rates of developmental screening? (2) Are there outcome differences for key subgroups of families? In particular, do outcomes differ for: (a) prenatally vs. postnatally enrolled mothers; (b) Hispanic vs. White/Caucasian mothers; (c) teenage vs. older mothers; (d) mothers with depressive symptomatology vs. non-depressed mothers; and (e) families with more vs. fewer	gagement with prenatal care services, single parent, depression (measured in mothers using PHQ-2), low education, drug abuse, troubled family relations. Families were deemed eligible if they scored positively for any two risk criteria or if there were substance abuse or depression issues. **Sample characteristics:** • Age - The average age of participants was 22.5 years. 31% were younger than 20, and 11% were younger than 18. (Measured at date of enrolment not date of interview. Demographics by group are not reported). • Sex - Not clear - all demographic data was measured using maternal status but it is unclear whether mothers or fathers participated in interviews. • Ethnicity - White 62%, Hispanic 24%, 'Other' (African American, American Indian/Alaska Native, Asian/Pacific Islander) 14% (Measured at date of enrolment not date of interview. Demographics by group are not reported). • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported.	pears to be significant in favour of intervention (reported as 'developmentally supportive activities - does not appear in table) - (F(1759)=5.162, p=0.023). This effect was more pronounced for non-depressed compared to depressed mothers, with non-depressed programme group mothers significantly more likely to engage in positive activities than non-depressed comparison group mothers ((F(1753)=4.15. p=0.042). Frequency of reading: Higher in intervention group (F(1759)=12.815, p=0.000). Breastfeeding: No significant difference Family functioning: No significant difference. **Effect sizes - Children and young people's health and wellbeing outcomes:** Child developmental status: Marginally significantly fewer HFO parents had been told that child had a developmental concern (OR=1.72, p=0.078). **Effect sizes - Caregiver/parent health and wellbeing outcomes:** Depressive symptomatology: No significant difference. However, for moderate and high risk families HFO families had significantly lower levels of depressive symptomatology (unclear what effect sizes are for this variable). **Narrative findings – Effectiveness:** Note that these findings relate to cross-sectional data gathered at 1 year - no baseline data gathered. Results suggest that HFO programme may help to reduce parenting-related stress (marginally lower stress levels in intervention compared to control), particularly in high risk group. Parenting stress has been shown to be associated	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
total risk factors' (p290). **Methodology:** RCT. **Country:** Not UK - USA. **Source of funding:** Government.	• Sexual orientation - Not reported. • Socioeconomic position - Single 80% 'Less than HS diploma/GED' 28% 'Both parents unemployed' 32% 'Difficulty paying expenses' 81% English as a first language 74% (p293) (Measured at date of enrolment not date of interview. Demographics by group are not reported). • Type of abuse – Not reported • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - Programme total: Not clear. Interviewed: n=401. • Intervention numbers - Programme total: Not clear. Interviewed: n=402. • Sample size - Programme total: 2664. Interviewed: 803. **Intervention category:** Home visiting. **Intervention:** The programme logic model focuses on three sets of activities, aiming to impact the long term outcomes of reduced child maltreatment and improved school readiness: 1. Increasing parents' knowledge of	with maltreatment. There were significantly more 'developmentally supportive' activities for HFO families. HFO families were also more likely to read to their infants, and more likely to have taken part in developmental screening. The results further suggests differential effectiveness for particular subgroups, including 'the program impact on parenting behaviours was larger for non-depressed mothers' (p296). There were also stronger programme impacts on parenting stress and depressive symptomatology for mothers with three of more risk factors. The highest risk HFO mothers were also significantly less likely to endorse the use of corporal punishment.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	child development to improve parenting knowledge and skills through parenting education, coaching and modelling. 2. Identification of family issues which may interfere with child development and parenting, such as depression. 3. Supporting healthy child development by promoting breastfeeding, developmental screening and promoting use of preventative health services. Home visiting provided by 1 of 7 programme sites. Little information on the visits themselves are provided, only that they are weekly for the first 6 months and can then be reduced depending on progress and level of need. Services provided for up to 3 years. Services provided by a 'trained home visitor'. **Comparison intervention:** Not clear - reports that '… comparison families were mailed a standard resource and referral information packet that is provided to all eligible families who are unable to be served by HFO' (p291). Not clear whether any additional services received after this. **Outcomes measured:** • Risk of abuse and neglect - Parent-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	ing stress measured using the parenting stress index (Abidin 1990), which has been shown to be associated with higher risk for maltreatment. • Quality of parenting and parent-child relationships - Self reported attitudes regarding corporal punishment were assessed using the Adult Adolescent Parenting Inventory, Corporal Punishment Subscale (AAPI-CP; Bavolek & Keene 2001). Parent–child interactions were assessed using the Parent–Child Activities Scale (PCAS; Love et al. 2002). The frequency with which parents read to their child was assessed by interview. Breastfeeding assessed via interview. Family functioning assessed using Family Functioning subscale of the Protective Factors Survey (Counts et al. 2010). • Children and young people's health and wellbeing outcomes - Child developmental status assessed via interview. • Caregiver/parent health and wellbeing outcomes - Depressive symptomatology assessed using a 3-item measure developed as part of the Pregnancy Risk Assessment Monitoring System (PRAMS 2008). **Follow-up:** Surveys were conducted		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	at children's one year birthday. NB no baseline measures were taken.		

9. Guterman NB, Tabone JK, Bryan GM et al. (2013) Examining the effectiveness of home-based parent aide services to reduce risk for physical child abuse and neglect: Six-month findings from a randomized clinical trial. Child Abuse and Neglect 37: 566–77

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: 'To examine the benefits of home-based paraprofessional parent aide services in reducing physical abuse and neglect in high risk patients' (p566). **Methodology:** RCT. **Country:** Not UK – USA. **Source of funding:** Government.	**Participants:** • Caregivers and families - Families (primarily mothers) deemed to be at high risk of abuse and/or neglect, measured by a referral from child protective services (CPS) or initial case assessment conducted by a programme staff member. Mothers had to be biological or adoptive mother of at least one child aged 12 years or younger. Mothers were ineligible if: Not fluent in English - Under 18 years old - Demonstrated a psychotic mental illness - Had a substance misuse problem for which they were not actively receiving treatment - Showed an IQ below 60. **Sample characteristics:** • Age - Mean age at 6-month follow-up: 29.7 (exposure group) and 29.5 (comparison group). • Sex - All participants appear to have been women.	**Effect sizes - Incidence of abuse and neglect:** Effect sizes calculated using difference in means at follow-up compared to baseline are reported as follows: case management only group - within group; case management plus parent aide services - within group; between groups. Differences in means that were found to be statistically significant using two-factor ANOVA are indicated using * p=0.10 and ** p=0.05. Child abuse and neglect measures: Incidence of abuse and neglect 1. Psychological aggression: -0.137; -.0259*; 0.122 2. Physical assault: -0.119; -0.304; 0.185 3. Household inadequacy: 0.084; 0.655**; -0.581* Maternal risk and protective factors 1. Parenting stress index: -0.117; -0.544**, 0.427 2. Maternal depression: -0.146; -0.313**; 0.167 3. Maternal anxiety: -0.214; -0.375**; 0.161 4. Maternal hostility: -0.049; -0.160; 0.111 5. Maternal drug use: 0.292; 0.219; 0.073 6. Male partner drug use: 0.074; -0,188; 0.262 7. Parental mastery: 0.042; 0.319**; -0.277 8. Maternal social support: 0.226; 0.051; 0.175. **Narrative findings – effectiveness:** The study found no significant differences between intervention (case management plus parent aide services) and control (case management only) groups on any outcomes,	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity score:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Ethnicity - Both exposure and comparison groups were just under 2/3 White (61.4% and 63.6% respectively), just under third Black (31.6% and 31.8% respectively and less than 1/10 other ethnicity (8.8% and 4.6%). • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Paper states that it is focusing on families at risk of physical abuse and neglect, although it is not clear if other types of abuse were also included. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - Baseline n=65 6-month post-test=44 Lost to follow-up n=21. • Intervention numbers - Baseline n=73 6-month post-test=57 Lost to follow-up n=16. • Sample size - Intervention group:	with the exception of the 'household inadequacy' measure intended as a proxy for neglect. The direction of the effect was that there was higher inadequacy in the intervention compared to the control group. This suggests that this group had actually deteriorated on this measure to a greater extent than the control group. The reasons for this are not explored in the Discussion section, which is disappointing. Statistically significant improvements within the intervention group (and not the control group) were observed for: - Psychological aggression (at p=0.10) - Parenting stress (at p=0.05) - Maternal depression (at p=0.05) - Maternal anxiety (at p = 0.05), and - Parental mastery (at p=0.05). However, the extent of improvement did not show a significant difference between the intervention and control groups. The study interprets this as suggestive of 'some modest positive benefits' (p575).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Baseline n=73 6-month post-test=57 Lost to follow-up n=16 Comparison group: Baseline n=65 6-month post-test=44 Lost to follow-up n=21. **Intervention category:** Parent aide services. **Intervention:** 'Parent aide' services in addition to case management services. Parent aides are 'paraprofessionals' receiving 12 hours of on-the-job training, followed by monthly training and regular supervision thereafter. Parent aides delivered services in the home specifically targeting: - child safety - parenting skill guidance - problem-solving support - improving parents' social support. Parent aides could visit up to twice per week, depending on assessed risk, need and assigned level of service. Frequency of visits begin with an intensive engagement phase, followed by a subsequent phase emphasising work on parent-child discipline and family communication, and later 'attenuating to focus on maintenance of gains and termination' (p570). Participants in the intervention condi-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	tion received an average of 17.45 contacts, totalling an average of 819.76 minutes over the six month period. **Comparison intervention:** Those assigned to 'Case management only' received an initial needs assessment, crisis intervention counselling where necessary and referrals for substance abuse, child care/respite and other community resources where necessary. They received limited (up to 2 per month) phone contacts or, if they did not have active phone lines, contacts in the home. Participants in the 'Case management only' condition received an average of 8.95 contacts, with a total length of contact averaging 207.19 minutes. **Outcomes measured:** • Incidence of abuse and neglect - 3 proxy measures were used, two self-report measures and 1 observational measure. Self-report measures: 1. The physical aggression, psychological aggression and neglect subscale items from the Parent-Child Conflict Tactics Scale (Straus et al. 1998) 2. Mother-Child Neglect Scale (Lounds et al. 2004) Observational measure: 1. Household adequacy scale within the Child Well-Being Scales (Magura & Moses 1986).		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Risk of abuse and neglect - 8 measures assessing potential risk and protective factors relating to parents (note these overlap with caregiver health and wellbeing outcomes). 1. Parenting Stress Index - Short Form (Abidin 1995) 2. Maternal depression measured by Brief Symptom Inventory (Derogatis 1975) 3. Maternal anxiety measured by Brief Symptom Inventory (Derogatis 1975) 4. Maternal hostility measured by Brief Symptom Inventory (Derogatis 1975) 5. Maternal drug use measured by Drug Use Screening Inventory (Tarter 1990) 6. Male partner drug use measured by Drug Use Screening Inventory (Tarter 1990) 7. Parental mastery measured by the Pearlin-Schooler Mastery scale, which reports parents' personal sense of control over life circumstances. 8. Maternal social support measured by the Multidimensional Scale of Perceived Social Support (Simet et al. 1988). **Follow-up:** Six months.		

10. Lam WKK, Fals-Stewart W, Kelley ML (2009) Parent training with behavioral couples therapy for fathers' alcohol abuse: effects on substance use, parental relationship, parenting, and CPS involvement. Child Maltreatment 14: 243–54

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The pilot study aimed to examine the '... effects of Parent Skills with Behavioral Couples Therapy (PSBCT) on substance use, parenting, and relationship conflict among fathers with alcohol use disorders' (p243) **Methodology:** RCT. **Country:** Not UK - USA. **Source of funding:** Government - National Institute on Alcohol Abuse and Alcoholism (R21AA013690).	**Participants:** • Children and young people. Children of heterosexual married or cohabiting men entering alcohol abuse treatment. If there was more than 1 child in the targeted age range, 1 was randomly selected to take part in the trial (i.e. interview - children were not directly involved in the treatment programme). • Caregivers and families. Heterosexual married or cohabiting couples in which the male was entering alcohol abuse treatment. **Sample characteristics:** • Age - Inclusion criteria: At least 18 years of age. Male partners' age (years): PSBCT 33.4 (5.1); BCT 34.6 (4.9); IBT 34.2 (4.4). Female partners' age (years): PSBCT 33.2 (5.4); BCT 32.8 (5.4); IBT 33.1 (5.2). Child's age: PSBCT 8.9 (2.1); BCT 9.0 (2.0); IBT 8.8 (2.2). • Sex – Male 50%, Female 50% (heterosexual couples) Male children (number (%)): PSBCT 6 (60); BCT 5 (50); IBT 5 (50). • Ethnicity - Male partners' race/ethnicity – White: PSBCT 7 (70); BCT 6 (60); IBT 6 (60). Male partners' race/ethnicity - African American:	**Effect sizes - Incidence of abuse and neglect:** Child maltreatment – parental self-report of involvement with services: Involvement with child protective services decreased in the treatment group across time (r>0.2). Small to moderate effect sizes compared with Individual-Based Treatment were observed at 6 and 12 months (r>0.2). Interparental conflict and violence - Timeline Followback Interview-Spousal Violence: Contrasts between groups were negligible (r<0.20). PSBCT and BCT showed within-group improvement over time (r>0.20). Interparental conflict and violence - Dyadic Adjustment Scale: This showed within-group improvement across all three conditions (r>0.30). Paired contrasts found medium effect sizes for PSBCT versus IBT at 6 and 12 months (r>0.20). **Effect sizes - Quality of parenting and parent-child relationships:** Parenting measured via the 'laxness' subscale of the Parenting Scale (self-reported by target child): Within-group scores for PSBCT showed medium effect sizes (r>0.30) from pre-treatment to each of the follow-up assessments. Paired contrasts between PSCBT and IBT were observed in the medium range (r>0.30) for mothers and small to medium range (r>0.20) for fathers. Parenting measured via the 'over-reactivity' subscale of the Parenting Scale (self-reported by target child): PSCBT showed reductions in over-reactivity at each follow-up assessment (r>0.20). Contrasts between PSCBT and BCT showed differences favouring PSCBT (r>0.20). Parenting measured using the Parental Monitoring Scale: PSCBT showed effects in medium range (r>0.30) from baseline to each	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity score:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	PSBCT 2 (20); BCT 3 (30); IBT 2 (20). Male partners' race/ethnicity – PSBCT Hispanic: 1 (10); BCT 0 (0); IBT 1 (10). Male partners' race/ethnicity – PSBCT Other: 0 (0); BCT 1 (10); IBT 1 (10). Female partners' race/ethnicity - White PSBCT: 7 (70); BCT 7 (70); ICT 6 (60). Female partners' race/ethnicity - African American: PSBCT 1 (10); BCT 1 (10); ICT 2 (20). Female partners' race/ethnicity - Hispanic PSBCT: 0 (0); BCT 1 (10); ICT 1 (10). Female partners' race/ethnicity - PSBCT Other: 2 (20); BCT 1 (10); ICT 1 (10). • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Heterosexual couples. • Socioeconomic position - Family income (annual in in US $1,000s): PSBCT 35.2 (15.6); BCT 34.0 (14.9); IBT 34.6 (15.3). • Type of abuse - No focus - the authors measured 'child maltreatment' through involvement with Child Protection Services. Interparental conflict and violence was also measured. • Looked after or adopted status - Not reported.	follow-up assessment. Paired contrasts with BCT and IBT showed medium effect sizes for mothers (r>0.30) and higher effect sizes for fathers (r>0.50) NB effect sizes differ in text compared to Table 4. **Effect sizes - Caregiver/parent health and wellbeing outcomes:** Substance use measured using the Timeline Followback Interview: All groups showed within-groups improvement in abstinence from alcohol. There were no significant differences between groups. **Narrative findings – Effectiveness:** The authors report that the results of the study indicate that 'couples therapy with parenting skills training (PSCBT) versus individual therapy showed effect size differences in parenting that approached the medium range across the 12-month period' (p.250). PSCBT was the only condition in which involvement with CPS services meaningfully decreased at each follow-up, and showed meaningful contrasts with IBT at 6 and 12 months. Changes in parenting resulting from PSBCT were greater in magnitude at each follow-up period relative to each of the other conditions. Fathers' reports of parenting resulted in stronger effects than mothers' reports.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** There were 3 treatment conditions giving a total sample size of 30 but numbers by group are not reported. **Intervention category:** Multi-component intervention. **Intervention:** Describe intervention Each condition comprised 24 sessions, with two 60-minute sessions per week for 12 weeks. The experimental condition was Parent Skills with Behavioral Couples Therapy (PSBCT): 12 sessions attended by both partners (6 sessions of core Behavioral Couples Therapy and six parent-skills training sessions.) The parent training sessions were developed using the programme designed by Forehand In 'Helping the Noncompliant Child' (Forehand and Long, 2002; McMahon and Forehand, 2003). Male participants also 12 attended individual cognitive behavioural therapy sessions. Masters' level therapists with experience in conducting Behavioral Couples Therapy and coping skills therapy for substance abuse were trained and		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	certified by the treatment developers to deliver all of the modules for each treatment. **Comparison intervention:** Behavioural Couples Therapy (BCT): 12 sessions attended by both partners. Aimed at improving communication, problem-solving and sobriety. Male participants also 12 attended individual cognitive behavioural therapy sessions. Individual-based treatment (IBT): 12 sessions attended by male participants only. These were modified coping skills sessions for alcoholism (Monti et al. 1989). Participants also attended 12 individual cognitive behavioural therapy sessions. **Outcomes measured:** • Incidence of abuse and neglect - Interparental conflict and violence (self-reported, calendar interview) measured using the Timeline Followback Interview–Spousal Violence (TLFB-SV; Fals-Stewart, Birchler, and Kelley 2003). Interparental conflict and violence was also measured using the Dyadic Adjustment Scale (Spanier 1976) - a measure of relationship satisfaction (self-reported by mother and father). Child maltreatment measured using parental self-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	report (dichotomous – 'Do you currently have an open case with CPS regarding the target child) of active involvement with Child Protection Services?' (p246). • Quality of parenting and parent-child relationships - Parenting measured (self-reported by target child) using two subscales of the Parenting Scale (Arnold, O'Leary, Wolff & Acker, 1993):- 1. 'Laxness' - permissive, inconsistent parenting. 2. 'Overreactivity' - harsh parenting. Parenting was also measured using the Parental Monitoring Scale (Bank et al. 1993) which self-reported. • Caregiver/parent health and wellbeing outcomes - Substance use measured using the Timeline Followback Interview (Sobell and Sobell 1996), a self-reported event history interview. At baseline assessments the measurement interval was the past 12 months. At subsequent assessment points the interval dated back to the date of the previous assessment. **Follow-up:** Baseline assessments conducted 1 week after admission to treatment programme. Follow-up took place on treatment completion, at 6 months and at 12 months.		

11. LeCroy CW and Krysik J (2011) Randomized trial of the healthy families Arizona home visiting program. Children and Youth Services Review 33: 1761–6

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To "... examine the effectiveness of home visiting as a means of improving parental, child, and maternal outcomes and preventing child abuse and neglect" (p1761). **Methodology:** RCT. **Country:** Not UK USA. **Source of funding:** Government - Arizona Department of Economic Security.	**Participants:** • Caregivers and families. Prenatal and new parent families screened and assessed as at risk of poor child and maternal outcomes and child abuse and neglect. Families were first screened using a 15-item risk checklist, (e.g. teen mother) and then assessed using a survey - a modified version of the Kempe Family Checklist. If a score of 25 or more was recorded the programme was offered to the family. **Sample characteristics:** • Age - Mothers average age at baseline: Intervention 23.5 years. Control 25.4 years. • Sex - Data appears to have been collected solely from mothers. • Ethnicity - White: Intervention 18.6%, control 23.7%. Hispanic: Intervention 64.9%, control 54.6%. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - AHCCCS health insurance (Arizona Medicaid programme): Intervention 95.7%, control 84.4% (significance of .01).	**Effect sizes - Incidence of abuse and neglect:** Effect sizes are not provided - The authors defined statistical significance at a level of p<.10. P values for contrasts between Healthy Families and control group are as follows. Violent behaviour – Disciplinary practices/aggressive discipline between groups: Six-month assessment not measured, 1-year assessment p=.10 Family violence: Six-month assessment p=0.15; one year assessment p=0.37. **Effect sizes - Quality of parenting and parent-child relationships:** The authors defined statistical significance at a level of p<.10. P values for contrasts between Healthy Families and control group are as follows. Parenting attitudes and practices – Inappropriate expectations: 6-month assessment p=0.10; 1-year assessment p=0.91 Lack of empathy: Six month assessment p=0.54; 1-year assessment p=0.91 Belief in corporal punishment: 6-month assessment p=0.12; 1-year assessment p=0.63 Reversing roles: 6-month assessment .32; one 1-assessment p=0.33 Oppressing child's independence: 6-month assessment p=0.06; 1-year assessment p=0.68 Safety practices: 6-month assessment p=0.04; 1-year assessment p=0.42 Mother's reading: 6-month assessment p=0.28; 1-year assessment p=0.85 **Effect sizes - Caregiver/parent health and wellbeing outcomes:** Effect sizes are not provided - only p values. The authors defined statistical significance at a level of p=<.10. P values for contrasts between	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity score:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Employment of mother: Intervention 17.7%, control 40.2% (significance of .000). Car ownership: Intervention 26.8%, control 53.6% (significance of .000). • Type of abuse - Involved with Child Protective Services as a parent: Intervention 24.7%, control 11.3% (significance of .01). History of childhood maltreatment - Neglected by caretakers: Intervention 24.7%, control 21.6%. Emotionally abused: Intervention 33.0%, control 19.6%. Sexually abused: Intervention 24.7%, control 21.6%. • Looked after or adopted status – Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - Baseline: n=97. Six month assessment: n=88. One year assessment: n=86. • Intervention numbers - Baseline: n=98. Six month assessment: n=92. One year assessment: n=85. • Sample size - n=195. Comparison: Baseline n=97, six month assessment n=88, one year assessment n=86. Intervention: Baseline n=98, 6-month assessment n=92, 1-year assessment n=85.	Healthy Families and control group are as follows. Parenting support - use of resources: 6-month assessment p=0.007; 1-year assessment p=0.001 Mental health and coping - Emotional loneliness: 6-month assessment p=0.34; 1-year assessment not measured. Pathways to goal: 6-month assessment p=0.12; 1-year assessment p=0.87 Alcohol use: 6-month assessment not measured; 1-year assessment p=0.04 Maternal outcomes - School or training: 6-month assessment not measured; 1-year assessment p=0.01 Using birth control: Six 6-assessment p=0 .61; 1-year assessment p=0.54. **Narrative findings – Effectiveness:** The study found (using a value of p=<.10. as definition of statistical significance - effect sizes are not provided.) significant differences across seven outcome measures. At six months there was a significant difference between groups on measures of inappropriate expectations, oppressing the child's independence, and safety practices. At the 1-year assessment there was a significant difference between groups on measures of aggressive discipline, alcohol use and maternal in engagement in education or training. At both the 6-month and 1-year assessment there was a significant difference between groups on the measure of parental use of resources. No significant differences were detected on measures of family violence, lack of empathy, belief in corporal punishment, reversing roles, mother's reading to the child, emotional loneliness, pathways to goal, and use of birth control at either the 6-month or 1-year assessment point.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Intervention category: Home visiting. **Intervention:** Home visiting services provided by the Healthy Families Arizona program. 'The overall goals of the program are to promote positive parenting, enhance child health and development, and prevent child abuse and neglect' (p1762). All home visitors at the site were female and had a degree or equivalent amount of experience. Home visitors aim to help new parents adapt to their new baby and address their life circumstances and needs. They can also help parents to access substance abuse, mental health and domestic violence services; encourage positive parenting behaviour, assess the developmental progress of children; improve safety in the home; give emotional support to parents; and help to secure a 'medical home' for the child. **Comparison intervention:** Control families received information from their child's developmental progress assessment. **Outcomes measured:** • Incidence of abuse and neglect - Two proxy measures were used to evaluate impact on child abuse and		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	neglect. The authors decided not to use official child protective reports due to risk of surveillance bias. Their primary outcome measure was mother's disciplinary practices and violence in the home. This was measured using a modified (shorter version – '… using the most serious indicators of abusive and neglectful behaviour …' p1763) of the Revised Parent-Child Conflict Tactics Scale (Straus, Hamby et al. 1998). Domestic/family violence was measured using an index created specifically for the study using common indicators similar to those used in the Conflict Tactics Scale (Straus et al. 1998). All measures were self-reported at interview. • Quality of parenting and parent-child relationships - 'Parenting attitudes and practices' were measured using the Adult-Adolescent Parenting Inventory-2 (Bavolek 1994). 'Parenting attitudes and practices' were also measured via safety in the home (e.g. car seat, poisons, etc. - true/false) and mother's reading to the child (estimated on a weekly basis). All measures were self-reported at interview. • Caregiver/parent health and wellbeing outcomes - 'Parenting support' measured use of resources such as		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	mental health and financial counselling and used a scale developed specifically for the study. The data was collected via self-report at interview. Mental health and coping was measured using 3 scales and was self-reported at interview. The first two were validated and established measures – the Emotional/Social Loneliness Inventory (DiTommaso and Spinner 1993) and a subscale of the Adult Hope Scale (Cramer and Dyrkacz 1998). Alcohol use was measured using a series of filter questions – yes/no – amount consumed in past two weeks. Maternal outcomes' were self-reported at interview (involvement in education or training and use of birth control). **Follow-up:** Six months and one year of age.		

12. Mejdoubi J, van den Heijkant SCCM, van Leerdam FJM et al. (2015) The effect of VoorZorg, the Dutch nurse-family partnership, on child maltreatment and development: a randomized controlled trial. PloS one 10: e0120182

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To investigate the effectiveness of VoorZorg, a Dutch adaptation of the	**Participants** Caregivers and families - 460 pregnant women with the following characteristics: <26 years of age,	**Effect sizes** Incidence of abuse and neglect From pregnancy to 3 years after birth, children in the intervention group were significantly less likely to	**Overall assessment of internal validity** ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Nurse-Family Partnership (NFP) in preventing child maltreatment. **Methodology:** RCT inc cluster. Single-blind (interviewers blinded but not participants or care-givers), parallel group, RCT randomising 460 first-time mothers from 20 regions in the Netherlands, aged less than 26 years and with low education attainment, to either a nurse home care visitation intervention - VoorZorg, (VZ)- or usual care. There was a two-stage selection criteria for participants which also identified women reporting 1 of 9 risk factors. Its primary outcome measure was child abuse reports and secondary outcome measure of child development, measured at 6 months, 18 months and 24 months of age, measured with the Home Observation Measurement of the	low educational level, first time pregnancy, maximum 28 weeks gestation, some understanding of the Dutch language. Women, from a larger group who met those 5 criteria, were also interviewed to assess whether they had at least 1 of the following 9 additional risk factors: being single; a history or current experience of domestic violence; psychosocial symptoms; unwanted pregnancy; financial problems; housing difficulties; no employment and/or education; alcohol and/or drug abuse. N=237 allocated to the intervention group (usual care + VoorZorg programme), n=223 allocated to the control group (usual care). Of the participants, n=168 of the intervention group and n=164 of the control group who were followed up 3 years after birth and analysed against Child Protection Services data, as these were participants from 8 CPS regions in the Netherlands that agreed to take part. **Sample characteristics** • Age - < 26 years of age. • Sex – Female. • Ethnicity - Control group: n=110 Dutch; n= 13 Turkish/Moroccan; n=58 Surinamese/Antillean; n=42 Other. Intervention group: n=115 Dutch; n=13 Turkish/Moroccan; n=64 Suri-	have had a Child Protective Services Report (RR=0.58, 95% CI 0.28 to 0.96). Subgroup analyses showed no significant differences when stratified by gender or ethnicity (no data reported). Quality of parenting and parent-child relationships From 6 months to 18 months there no significant difference between groups for IT-HOME scores, although they increased in both. (6 months MD 0.4; CI 95% -2.75 to 2.04; 18 months MD 0.80; CI 95% -1.30 to 2.91). At 24 months the IT-HOME score for the intervention group was significantly higher than in the control group (MD 1.98; CI 95% 0.16 to 3.80) However, after mixed model analyses (correcting for age of mother, ethnicity and several risk factors) there was no significance difference between groups in total IT-HOME scores over time (MD 1.12; CI 95% -0.59 to 2.83) Children and young people's health and wellbeing outcomes The number of children with internalising behaviour (measured by the CBCL at 24 months) was significantly lower in the intervention group than in the control group (RR 0.56; CI 95% 0.24 to 0.94, ARD 0.14). There was no significant difference between groups in the number of children with externalising behaviour (RR 0.71, CI 95% 0.34 to 1.09, ARD 0.10). Caregiver/parent health and wellbeing outcomes This outcome was measured by a different study.	**Overall assessment of external validity** ++ Only country is not exactly relevant. **Overall validity score** + Relatively high attrition rate for the study (32.8% for whole study sample).

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Environment (IT-HOME) and the Child Behaviour Checklist 1.5–5 years (CBCL 1.5–5). There were other outcomes reported in a different study of the same trial which looked at maternal smoking during pregnancy, birthweight, breastfeeding and incidence of IPV. Poisson regression models were used to analyse CPS reports and CBCL/1.5-5 scores, while relative risks, absolute risk differences and CIs were calculated using a Poisson log-linear model according to Zou. Multiple imputation models were used to account for issuing CBCL data at 24 months and these were validated with IBM SPSS generated sensitivity analysis which generated 50 imputed datasets as recommended. IT-HOME scores were	namese/Antillean; n=45 Other. Ethnicity based on self-reporting, classified as certain ethnicity if 1 or more of her biological parents was born outside the Netherlands. • Religion/belief - Not reported. • Disability - Not reported • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position – Low educational attainment. • Type of abuse – Not reported. • Looked after or adopted status – Not applicable. • Unaccompanied asylum seeking, refugee or trafficked children - Not applicable. **Sample size** Comparison numbers - n=223. Intervention numbers - n=237. Sample size - 460 in total n=237. intervention n=223 control. **Intervention** Intervention category – Family Nurse Partnership. Dutch adaptation of Nurse-Family Partnership (NFP). Describe intervention - VoorZorg (VZ) is a programme of home visits carried out by nurses trained and experienced in the VZ method. Visits in-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
analysed with multiple linear regression and then mixed model analyses to measure the longitudinal relationship between VZ and IT-HOME scores. All analyses were adjusted for any possible confounders/effect modifiers. Attrition analysis was also conducted. **Country:** Not UK. Netherlands. **Source of funding:** Government. Netherlands Organisation for Health Research and Development.	volve conversations relating to pregnancy and child development. They also involve health education, parenting skills lessons and building self-efficacy (in order to reduce the risk factors for child maltreatment). Nurses also work with mothers to look at how local social and community resources can be used to best benefit them and their child. Nurses and mothers communicate by SMS, phone and social media. Delivered by Nurses trained prior to delivery, supervised throughout and supported by biannual one-day national training sessions (involving role play) and peer-observation during visits. Delivered to Young mothers from pregnancy through to 36 months after the birth of their child. Duration, frequency, intensity, etc. 10 visits during pregnancy, 20 during the first year of life of the child, 20 visits during the second year of the life of the child. Key components and objectives of intervention The goals, procedures and content of visits are detailed for nurses in their		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	manuals, which have been translated from Nurse-Family Partnership manuals. Conversations during visits cover 6 domains relating to the relevant stage of pregnancy or child development. The purpose of these conversations, alongside the content looking at health education, parenting skills and self-efficacy is to reduce the risk factors for child maltreatment. The varied communication methods are designed to support the development of enduring relationships of trust between the nurse and the mother. Location/place of delivery Home of the mother. Describe comparison intervention Usual care. **Outcomes measured** Incidence of abuse and neglect. Child neglect comprising: 1. Affection neglect - measured using Coder Impressions Inventory (Dishion et al. 2004) at ages 2 and 4 2. Monitoring neglect - measured using Code Impressions Inventory (Dishion et al. 2004) at ages 2 and 4 3. Caregiving neglect - measured using Home Observation for Measurement of Environment (Caldwell and Bradley 2003).		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Quality of parenting and parent-child relationships. Observed dyadic positive engagement - measured using the Relationship Affect Coding System (Peterson et al. 2008) at ages 2 and 3. **Follow-up** Assessments were conducted at baseline, age 2, age 3, age 4 and age 5). However, only data from ages 2 and 3 were used in relation to dyadic positive engagement, and only data from ages 2 and 4 were used in relation to neglect. **Costs** No.		

13. Nelson HD, Selph F, Bougatsos C et al. (2013) Behavioral Interventions and Counseling to Prevent Child Abuse and Neglect: Systematic Review to Update the U.S. Preventive Services Task Force Recommendation. Rockville: Agency for Healthcare Research and Quality

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To examine: '1. For children without obvious signs and symptoms of abuse or neglect, but potentially at increased risk, how well do behavioral interventions and counseling initiated in primary care settings reduce exposure to abuse	**Participants:** • Children and young people. Children aged 0-18 years who are 'asymptomatic' - that is, not showing any signs of abuse or neglect. All studies referred to children <5 years. • Caregivers and families. Pregnant women and mothers. Studies included where risk factors for child abuse and neglect were identified.	**Effect sizes - Incidence of abuse and neglect: Clinic-based intervention (1 trial)** Families in the intervention group had fewer CPS reports than usual care group up to 44 months after the intervention (13% vs. 19%, p=0.03). Parents in the intervention group reported fewer episodes of severe or very severe physical assault than usual care parents (p=0.04).	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** + **Overall validity score:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
or neglect, physical or mental harms, or mortality? 2. What are the adverse effects of behavioral interventions and counseling to reduce harm from abuse and neglect?' (p6). **Methodology:** Systematic review - Systematic review of 11 RCTs (1 clinic-based and 10 home visitation programmes). Results are also compared to 8 RCTs of home visitation programmes reviewed in a previous report (US Preventive Services Task Force 2004). **Country:** Range of countries. Review conducted in US, included studies mainly US, also UK, Canada, Australia and NZ. **Source of funding:** Government - US Department of Health and Human Services.	**Sample characteristics:** • Age - Children 0-5 years; pregnant women aged 17-24 years; young mothers mean age 24 years. • Sex - Children: not reported. • Ethnicity - 40-80% Black, Latino families; also some from minority groups (Maori in NZ); Native Hawaiians, Pacific islanders, Alaskan natives, Asians and Filipinos. • Religion/belief - not reported. • Disability - not reported. • Long term health condition - not reported. • Sexual orientation - not reported. • Socioeconomic position - included some families below poverty levels. • Type of abuse - not reported. • Looked after or adopted status - not reported. • Unaccompanied asylum seeking, refugee or trafficked children - not reported. **Sample size:** • Comparison numbers - available only for some studies. • Intervention number - available only for some studies. • Sample size - 6,608 families (parents and children). • Systematic reviews (number of stud-	Parents in intervention group also showed fewer instances of non-adherence to medical care (p=0.05) and fewer delays in immunisations (p=0.002). **Home visiting interventions (10 trials)** 1. CPS involvement (measured in 6 trials): No trials found differences in rates of CPS reports while studies were ongoing. There was a significant intervention effect in one trial relating to decreased CPS involvement at 3 years after enrolment. (Lowell et al. 2011) 14% vs. 31%; OR, 2.1 (95% CI, 1.1–4.4); p<0.05 There was no significant differences in total CPS reports after either 1 or 3 years of follow-up (3 trials). Three other trials did not present any data on CPS due to very low rates of family participation. 2. Legal removal of the child from home (measured in 2 trials) There was no significant intervention effects between intervention and control at 36 months follow-up. 3. Self-reports of child abuse and neglect (measured in 5 trials). One trial found a significant intervention effect in self-reported severe physical assault at 36 months (Fergusson et al. 2005). 4.4% vs. 11.7%; OR 0.35 (95% CI 0.15 - 0.80); p<0.01. A second study found significant differences in favour of the intervention in very serious abuse (year 1 only - 0.01 vs 0.08; p=0.04) and serious physical abuse (year 2 only - 0.01 vs 0.04; p=0.03) (DuMont et al. 2008). Three further trials found no significant differences in rates of self-reported abuse.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	ies) - 11 trials* of interventions (conducted between 1980 and 2011), one conducted in the UK, one in New Zealand and 17 in the US) * this also included 8 trials which were extension of previous trials. **Intervention category:** • Home visiting - see information below on intervention for further details. • Parenting programmes - most home visiting programmes involved some parenting element (see information below on intervention for further details.) **Intervention:** 19 interventions (including 8 trials which were extension of previous trials). • 1 clinic-based: University-based paediatric primary care resident continuity clinic serving a low-income urban population. Intervention involved – '… 1) specially trained residents, including handouts for doctors and patients; 2) administration of the Parent Screening Questionnaire; 3) a social worker' (Dubowitz 2009) (p90) • Home visitation programmes 'Home visits began either before or after birth and continued for 3 to 36 months after birth. The intervention	**Effect sizes - Children and young people's health and wellbeing outcomes:** 1.Child mortality (measured in one 1 trial): There was no significant intervention effect on child mortality (Olds et al. 2007). 2. Emergency visits (measured in 3 trials). Reduced hospital visits for injuries and ingestions was reported in one trial (Fergusson et al. 2005). 17.5% vs. 26.3%; OR 0.59 (95% CI 0.36-0.98); p<0.05. One trial found a significantly greater proportion of mothers in the intervention group had never used the emergency room for child health problems (Koniak Griffin et al. 2003). 36% vs 11 %; p<0.05. 3. Hospitalisations (measured in 6 trials) There was no significant intervention effects in hospitalisation (5 trials) and one trial reported fewer hospitalisation episodes at 12 and 24 months (Koniak Griffin et al. 2003). 19 vs 36, p>0.01. 4. Adherence with child immunisations and well-child visits (measured in 3 trials). In one trial, home-visited children received immunisations at an earlier age than children in the control group, meaning that there were significant differences in rates of immunisation at 9, but not 12 months (El-Mohandes et al. 2003). 2.20 vs. 1.64; p=0.0125. Two other trials showed no difference in rates of immunisation at 2 and 3 years. **Narrative findings – effectiveness** **Clinic-based intervention (1 trial)**	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	was provided by either a paraprofessional, such as a lay person who had participated in a 9-week training course (nine trials), or a professional, typically a nurse (five trials)' (p10). • Home visiting example 1: Memphis Study (Olds 2007) 1) Transportation to clinic 2) Same as group 1 plus developmental screening and referral services at 6, 12, and 24 months 3) Same as groups 1 and 2 plus 3 intensive home visitations 4) Same as groups 1, 2, and 3 plus intensive home visitation services through age 2 years (App B1). • Home visiting example 2:Family Partnership Model (1 trial: Barlow 2007) 1) Control 2) 18 months of weekly visits from a heath visitor trained in understanding the processes of helping, skills of relating to parents effectively, and methods of promoting parent-infant interaction (App B1). • Home visiting example 3: Healthy Families Alaska (1 trial: Duggan 2007) Home visiting for 3–5 years, offered weekly for the first 6–9 months; families are promoted to service levels with less frequent visits as family functioning improves. Home visitation includes information, referrals, preparation of parents for	Families in the intervention group had fewer CPS reports than usual care group up to 44 months after intervention. There were also fewer episodes of severe or very severe physical assault in the intervention group, fewer instances of nonadherence to medical care and fewer delays in immunisations. **Home visiting interventions (10 trials)** 1. Child mortality (measured in one 1 trial): There was no significant intervention effect on child mortality (Olds et al. 2007). 2. CPS involvement (measured in 6 trials): No trials found differences in rates of CPS reports while studies were ongoing. There was a significant intervention effect in one trial relating to decreased CPS involvement at 3 years after enrolment. (Lowell et al. 2011) Subgroup analysis in 1 trial at 2-year follow-up found that poor, high-risk teenage mothers who were visited by nurses were less likely to commit acts of confirmed child abuse and neglect compared with those without visits, but the results at 3- and 4-year follow-up showed no differences. At the 15-year follow-up, children in the nurse-visited group were less likely to be involved in substantiated CPS reports. Nurse-visited mothers less likely to be a substantiated perpetrator of child abuse over the same 15-year period. There was no significant differences in total CPS reports after either 1 or 3 years of follow-up (3 trials). Three other trials did not present any data on CPS due to very low rates of family participation. 3. Legal removal of the child from home (measured in 2 trials) There was no significant intervention effects	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	developmental milestones, promotion of child environmental safety, and encouragement of positive parent-child interaction (App B1). • Home visiting example 4: Hawaii Healthy Start Program (1 trial: Duggan 2004) Home visits for 3–5 years by trained paraprofessionals to provide assistance, education, and services; model effective parent-child interaction; ensure child has medical home. Level 1: visited weekly; Level 2: biweekly; Level 3: monthly; Level 4: quarterly, with explicit criteria for promotion; intervention was for 1, 2, or 3 years (App B1). • Home visiting example 5: Healthy Families New York (1 trial: DuMont 2008) Home visits by trained paraprofessionals to provide assistance, education, and services; model effective parent-child interaction; ensure child has medical home (App B1). • Home visiting example 6: Early Start Program (1 trial: Fergusson 2005) Early Start Program assesses needs and resources, encourages positive partnership, provides support and problem solving (App B1). • Home visiting example 7: Cognitive-based extension of the Healthy Start Program (HSP) home visitation programme (Bugental 2009) Cognitive-	between intervention and control at 36 months follow-up. 4. Emergency visits (measured in 3 trials). Reduced hospital visits for injuries and ingestions was reported in one trial (Fergusson et al. 2005). There was no significant difference in rates of visits between intervention and control in the other 2 trials. However, one trial found a significantly greater proportion of mothers in the intervention group had never used the emergency room for child health problems compared with those in control group (Koniak Griffin et al. 2003). 5. Hospitalisations (measured in 6 trials) There was no significant intervention effects in hospitalisation (5 trials) and one trial reported a fewer hospitalisation episodes at 12 and 24 months (Koniak Griffin et al. 2003). 6. Adherence with child immunisations and well-child visits (measured in 3 trials). In one trial, home-visited children received immunisations at an earlier age than children in the control group, meaning that there were significant differences in rates of immunisation at 9, but not 12 months (El-Mohandes et al. 2003). Two other trials showed no difference in rates of immunisation at 2 and 3 years.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	based extension of the HSP home visitation vs. standard HSP home visitation vs. control condition. The additional cognitive appraisal component was designed to enhance parents' perceptions of power and competence, and included reframing in primary and secondary appraisals. No control group in Bugental 2009 (App B1). • Home visiting example 8: Child First (1 trial: Lowell 2011). Each family assigned a clinical team, consisting of a master's level developmental/mental health clinician and an associate's or bachelor's level care coordinator/case manager. Engagement and building trust were fundamental goals of Child First. No set curriculum (App B1). **Intervention category:** • Clinic-based. • Home visiting. **Comparison intervention:** Most control groups received standard care or usual service, or no home visits. See information above regarding interventions. **Outcomes measured:** • Incidence of abuse and neglect - All		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	studies but one looked at either incidence or risk of child abuse and neglect. Incidence measured via: CPS reports/involvement (6 trials), legal removal of child from home (2 trials), confirmed reports of abuse/neglect, self-reports of abusive behaviour using the Parent-Child subscale of the Conflict Tactics Scale (5 trials), instances of severe or very severe physical assault, new cases of child abuse/neglect, frequency of harsh parenting or physical abuse or spanking/slapping. • Risk of abuse and neglect - Beliefs about child-rearing associated with child abuse and neglect. • Children and young people's health and wellbeing outcomes - number of well-child visits; healthcare utilisation; delayed immunisations; non-adherence to medical care; emergency department visits. **Follow-up:** Duration of intervention in included trials ranged from 6 months to age 15 (children). Majority follow-up under 3 years.		

14. Peacock S, Konrad S, Watson E et al. (2013) Effectiveness of home visiting programs on child outcomes: A systematic review. BMC Public Health 13: 17

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: 'To systematically review the effectiveness of paraprofessional home-visiting programs on developmental and health outcomes of young children from disadvantaged families' (p1). The specific research question was: 'What is the effectiveness of paraprofessional HV programs in producing positive developmental and health outcomes in children from birth to six years of age living in socially high-risk families?' (p2). Six included studies specifically report child abuse and neglect outcomes. **Methodology:** Systematic review. 21 RCTs included, only studies from the US, Ireland and UK (n=17) to be included in our analyses.	**Participants:** • Children and young people. Socially high risk families - drug users, teenage mothers, low income, single mothers, poor neighbourhoods, children at poor school readiness, undernourished children. • Caregivers and families. Socially high risk families - drug users, teenage mothers, low income, single mothers, poor neighbourhoods, children at poor school readiness, undernourished children. **Sample characteristics:** • Age - Children aged 0–6 years. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position socially high risk families. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. Unaccompanied asylum seeking, refugee or trafficked children - Not reported.	**Effect sizes - Incidence of abuse and neglect:** A - Maltreatment Six studies measured maltreatment outcomes. Three showed no significant differences between intervention and control (Barth 1991; Duggan et al. 2004a; Duggan et al. 2004b) One study (Bugental et al. 2002) found that enhanced group had: less harsh parenting ($p=0.05$) less likelihood of physical abuse ($p<0.05$) and less likelihood of slapping/spanking children ($p<0.05$) compared to other groups. Duggan et al.'s evaluation of Healthy Families Alaska found decreased rates of substantiated maltreatment among those receiving home visitation services ($p<0.05$). DuMont et al.'s evaluation of Healthy Families New York found no overall programme effects, but did find differences for a 'prevention' subgroup of young first-time mothers, who were less likely to report minor physical aggression in the previous year ($p=0.02$) and harsh parenting behaviours in the previous week ($p=0.02$). The 'psychologically vulnerable' subgroup (older mothers with higher rate of CPS reports) were less likely to report acts of serious abuse or neglect compared to the control group at year two ($p<0.05$). **Effect sizes – Child health and wellbeing** B - Developmental delay B1 Psychomotor and cognitive development: Measured in 6 studies conducted in within scope countries. Three no significant impact. Three showed intervention	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** + **Overall validity score:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Country: Range of countries. US (14), UK (1), Ireland (2), Chile (1), Jamaica (1), South Africa (1), Bangladesh (1). **Source of funding:** Not reported. The authors have no financial relationships relevant to this article to disclose. The authors have no conflict of interest to disclose.	**Sample size:** • Comparison numbers - Not reported. • Intervention numbers - Not reported. • Sample size - 17 RCTs from US, UK and Ireland (N= 5447 families + children); 6 RCTs with child maltreatment outcomes (n=2552 families +children). • Systematic reviews (number of studies) - 17/21 RCTs from within scope countries (6 RCTs with child maltreatment outcomes also included in the review by Nelson 2013: Barth 1991; Duggan 2004a; Duggan 2004b; Duggan 2009; Bugental 2002; DuMont 2008.) **Intervention category:** Home visiting - All for at-risk families: Child Parent Enrichment Project; Healthy Families Alaska; Healthy Start Program; Healthy Families New York; Community Mothers' Programme; Seattle Birth to 3 Program; Hawaii Healthy Start Program; Home for Parents of Preschool Youngsters. **Intervention:** Paraprofessional home-visiting programme: (17 studies conducted in USA, Ireland and UK) Programmes with maltreatment outcomes:	effects (on BSID, development quotient and developmental stimulation). B2 Child behaviour: Measured in 1 study. Showed significant intervention effect on internalising/externalising behaviour. B3 Language development: Measured in 4 studies. Two no significant impact. Two intervention effects, though one was on expressive language only. C - Health assessments C1 Physical growth - Measured in 5 studies. Three no significant impact. One showed impact on birth weight, one on rehabilitating malnutrition. C2 - Hospitalisations, illness or injuries - Measured in 5 studies. Three no significant impact. Two positive impact on health outcomes. C3 Immunisations - Measured in 1 study. Impact shown in intervention group. **Narrative findings – Effectiveness:** 6 studies measured maltreatment outcomes. Three showed no significant differences between intervention and control (Barth 1991; Duggan et al. 2004a; Duggan et al. 2004b) One study (Bugental et al. 2002) found that enhanced group had: less harsh parenting less likelihood of physical abuse and less likelihood of slapping/spanking children compared to other groups. Duggan et al.'s evaluation of Healthy Families Alaska found decreased rates of substantiated maltreatment among those receiving home visitation services	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	1. Child-Parent Enrichment Project, pregnant women received, on average, 11 home visits over a 6 month period (Barth 1991). 2. Home visitation and cognitive change component (enhanced group) vs Healthy Start Model (un-enhanced group), pregnant women received, on average, 12 home visits over a year (Bugental et al. 2002). 3. Healthy Start Program, at-risk families received home visits (frequency varied) for up to 3-5 years (Duggan, Fuddy et al. 2004; Duggan, MacFarlane et al. 2004). 4. Healthy Families Alaska, at-risk families received home visits (frequency varied) for up to 3 years (Duggan 2009). 5. Healthy Families New York, at-risk families received home visits weekly to biweekly for up to 5 years (DuMont 2008). Programmes with outcomes on physical growth (weight + height): 1. No Name - Children with non-organic failure to thrive received home visits weekly for up to 1 year (Black 1995). 2. Healthy Families New York, At-risk adolescent mothers received home visits bi-weekly from pregnancy (Lee 2009). 3. No Name - At-risk pregnant women	($p<0.05$). DuMont et al.'s evaluation of Healthy Families New York found no overall programme effects, but did find differences for a 'prevention' subgroup of young first-time mothers, who were less likely to report minor physical aggression in the previous year ($p=0.02$) and harsh parenting behaviours in the previous week ($p=0.02$). The 'psychologically vulnerable' subgroup (older mothers with higher rate of CPS reports) were less likely to report acts of serious abuse or neglect compared to the control group at year 2 ($p<0.05$). Ten studies conducted within scope countries examined different aspects of developmental delay. The evidence for impact was equivocal, with roughly equal balance of significant and non-significant findings across psychomotor and cognitive, child behaviour and language development outcomes. Nine studies conducted within scope countries examined health assessments. Again, evidence was equivocal, with a roughly equal balance of significant and non-significant findings across physical growth, hospitalisations, illness and injury. One study relating to immunisations had a significant impact. Conclusion: Overall, home visitation programmes '… that utilize paraprofessionals often do not have significant effects on disadvantaged families, but show promise. However young children in these programs show modest improvements in some circumstances' (p13).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	received home visits (frequency not stated) from pregnancy (McLaughlin 1992). 4. No Name – At-risk children (from poor school readiness) received home visits monthly for up to 1 year (Scheiwe 2010). Programmes with outcomes on hospitalisations, illness or injuries: 1. Home visitation and a cognitive change component (enhanced group) vs Healthy Start Model (un-enhanced group), pregnant women received, on average, 12 home visits over a year (Bugental 2002). 2. Healthy Families Alaska, families received home visits (frequency varied) up to 2 years (Caldera 2007). 3. Healthy Start Program, at -risk families received home visits (frequency varied) for up to 3-5 years (Duggan 2004). 4. Community Mothers' Programme, first time mothers with children received home visits monthly up to 1 year (Johnson 1993). 5. No Name, at risk children (from poor school readiness) received home visits monthly for up to 1 year (Scheiwe 2010). Programmes with outcomes on up-to-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	date immunisations: Community Mothers' Programme, first time mothers with children received home visits monthly up to 1 year (Johnson 1993). **Comparison intervention:** All control group participants received the usual services offered in their community (p 3). **Outcomes measured:** Incidence of abuse and neglect - Outcome measures not clearly described. Appears to be: Self-reported abuse and neglect (Barth 1991) Self-reported harsh parenting, measured using Conflict Tactics Scale (Bugental et al. 2002) Child Protective Service reports (Duggan et al. 2009) Minor physical aggression and harsh parenting (DuMont et al. 2008). Children and young people's health and wellbeing outcomes - Report on health assessment: physical growth; up-to-date immunisations; hospitalisations, illness and injuries (Table 5). Report on developmental delays: psycho-motor and cognitive development; child behaviour; language development. **Follow-up:** Intervention durations ranged from pregnancy to 5 years.		

15. Pereira M, Negrão M, Soares I et al. (2015) Decreasing harsh discipline in mothers at risk for maltreatment: A randomized control trial. Infant Mental Health Journal 35: 604–13

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To assess the effectiveness of the Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD) at improving maternal sensitivity and reducing harsh discipline among 43 severely deprived Portuguese mothers of children aged 1–4 years, for whom there were concerns raised about the caregiving environment. **Methodology:** RCT inc cluster. Eligible families (exhibiting 1 of 23 risk factors relating to family relations or parenting quality e.g. neglect of child's health/emotional/cognitive needs, coercive discipline, lack of parental flexibility/self-	**Participants** Children and young people - Children aged 1- to 4-years living with their biological mother Caregivers and families - 43 Mothers of 1- to 4-year olds, known to health and social care services in Northern region of Portugal, for whom there are concerns about the caregiving environment in which the child is being raised. **Sample characteristics** • Age - Mean age of mother at baseline=29.86 (SD=6.22, range=18–46 years). (Intervention group - m=30.43, Control group m=29.32) Mean age of children at baseline=28.44 months (SD=10.38, range=12-48 months). (Intervention group 29.33 months (SD 9.71), control group m=27.59 months (SD 11.15). • Sex – Child gender Control group n=11 boys, n=11 girls Intervention group n=11 boys, n=10 girls. • Ethnicity - Ethnic minorities excluded because 'they would probably require adaptations to the standardized VIPP-SD used' (p 607). Only	**Effect sizes** Incidence of abuse and neglect From pregnancy to three years after birth, children in the intervention group were significantly less likely to have had a Child Protective Services Report (RR=0.58, 95% CI 0.28 to 0.96). Subgroup analyses showed no significant differences when stratified by gender or ethnicity (no data reported). Quality of parenting and parent-child relationships From 6 months to 18 months there no significant difference between groups for IT-HOME scores, although they increased in both (6 months MD 0.4; CI 95% -2.75 to 2.04; 18 months MD 0.80; CI 95% -1.30 to 2.91). At 24 months the IT-HOME score for the intervention group was significantly higher than in the control group (MD 1.98; CI 95% 0.16 to 3.80). However, after mixed model analyses (correcting for age of mother, ethnicity and several risk factors) there was no significance difference between groups in total IT-HOME scores over time (MD 1.12; CI 95% -0.59 to 2.83). Children and young people's health and wellbeing outcomes The number of children with internalising behaviour (measured by the CBCL at 24 months) was significantly lower in the intervention group than in the control group (RR 0.56; CI 95% 0.24 to 0.94, ARD 0.14). There was no significant difference between groups in	**Overall assessment of internal validity** - **Overall assessment of external validity** + Entirely relevant but not UK. **Overall validity score** - Small sample, no blinding, lack of information on whether study was sufficiently powered, also high levels of attrition as study progressed and unclear what impact this had on representativeness of sample.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
control/self-competence or domestic violence) were assessed using a Portuguese short version of the Family Risks and Strengths Profile (PRF) by social and health agency staff in the Northern region of Portugal. From an initial referral of n=156, the final eligible sample (of Portuguese children living with their biological mother) was n=44, with one 1 case excluded from analysis due to incomplete baseline data. Of the participating mothers, they were assessed at home at baseline (to encourage continued participation) in 2 pretest sessions within 2 weeks of each other, the first session was used to present and explain the research procedures and to collect written informed consent and self-reported questionnaire data from the mothers,	Portuguese ethnicity mothers included. • Religion/belief - Not reported. • Disability - Severe medical conditions for mother or child (severe physical impairment, intellectual deficit or psychosis) led to exclusion from eligibility for adaptation reasons. • Long term health condition - Severe medical conditions for mother or child (severe physical impairment, intellectual deficit or psychosis) led to exclusion from eligibility for adaptation reasons. • Sexual orientation - Not reported. • Socioeconomic position – Low family education attainment - 70.4% mothers, 86.4% of fathers of children had not completed Portuguese mandatory education (9 years). High levels of unemployment among parents (70.5% mothers, 50% fathers). Majority received welfare assistance (79.5%) and lived in poor housing conditions (77.3%). • Type of abuse – Eligible families were those referred to services who were scored to have 1 of 23 risk factors on the PRF relating to family relations quality or parenting quality: neglect of child's health/emotional/cognitive needs, coercive discipline, lack of parental flexibility/self-control/self-competence or domestic violence. Mean number of risk factors	the number of children with externalising behaviour (RR 0.71, CI 95% 0.34 to 1.09, ARD 0.10). Caregiver/parent health and wellbeing outcomes This outcome was measured by a different study.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
the second to gather over a 1 hr session videotaped observation of several mother-child interactions. The mothers were then randomised by researchers, stratified by child's age, gender, and temperament (coded using the Difficult Temperament subscale from the Infant characteristics Questionnaire (Bates et al. 1979)) to receive either the VIPP-SD intervention via 6 home visits or a control intervention of 6 telephone calls at the same intervals as the VIPP-SD sessions. A posttest of the same procedures and assessments as the pretest was completed with families 1 month after the last visit or phone call. An outcome of harsh discipline was measured pre- and post-test using observations of two tasks: 1. the clean-up task where the mother	in intervention group (n=21) 6.00 (SD 3.46). Mean number of risk factors in control group (n=22) 5.00 (SD 4.63). Harsh discipline outcome was calculated from coding for physical and verbal discipline and level of psychological control in content of maternal statements. Mean score at baseline - control group -.41 (SD 2.08), intervention group .34 (SD 2.22). • Looked after or adopted status – Only children living with biological mother eligible. • Unaccompanied asylum seeking, refugee or trafficked children - Only children living with biological mother eligible. **Sample size** Comparison numbers - n=22. Intervention numbers - n=21. Sample size - n=43, n=22 in control group, n=21 in intervention group. The pre-test data for one participant in original (n=22) intervention group was not available and they were therefore excluded from analysis. **Intervention** Intervention category – Parenting intervention. Describe intervention - VIPP-SD is a home-based intervention in which mother-child dyad interactions are		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
was instructed that the child should clean up as much as possible but that they should help and support them as they usually would, with the task ending after all toys were placed back in a box or 4 minutes elapsed or 2. the don't-touch task whereby the mother was instructed to place all toys from a box in front of their child, not allow them to touch them, and then allow them to play with the least-attractive toy after 2 minutes, with the task ending after a further 2 minutes. Physical, verbal and psychological control were coded on a scale of 1 (not present) to 5 (much frequent/continuous). Harsh physical discipline was coded when mothers showed unnecessary levels of physical force which had clear impact on child. Harsh verbal discipline was coded	videotaped. The tapes are then shown to the mothers in the next session (following further recorded observation) and discussed, with feedback provided on themes of parental sensitivity and discipline, as well as child development. Topics will vary in emphasis according to needs of mother-child dyad, as relating to their profile, built up by intervener from pretest videotaped observation session, a profile that is evaluated and reformed after each intervention session. Delivered by Four female intervenors, 3 who completed a 1 week training in VIPP-SD at Leiden university (including supervised working with a pilot family) and 1 other who was trained in Portugal by two of the study authors Pereira and Negrao. Delivered to Mothers of 1–4 year old children, already known to agencies to be risk of maltreating child, although in the intervention group fathers were asked to participate in the final two sessions (only n=2 took part). Duration, frequency, intensity, etc. 6 1 hr home-based intervention sessions, the first four scheduled at 2		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
when mother displayed high levels of irritation or anger in her tone of voice. Psychological control was coded according to the extent maternal statements made children feel guilty, ashamed or responsible for mishaps, when mothers showed disregard for child's feelings; withheld affection or displayed inconsistent emotional behaviour. Total scores were calculated by standardising and summing the 3 subscale scores for the two observed tasks. Parenting stress was measured using a self-reported Daily Hassles Questionnaire (Kanner et al. 1981) at pre- and post- test. Random assignment was checked using t tests and chi-square tests for demographic and pretest variables, with analyses adjusted for three harsh discipline outlier variables. Intervention	week intervals, with two further booster sessions a month apart. Key components and objectives of intervention First 4 sessions, designed to enhance maternal sensitivity and promote positive discipline addressed the following topics: 1. Difference between attachment and exploration behaviour, using distraction and induction as disciplinary tactics. 2. 'Speaking for the child' - helping mother to understand child cues and communication, using positive reinforcement as a disciplinary strategy. 3. 'Chain of sensitivity': understanding the child signal-mother recognition-mother interpretation-mother response-child response, using sensitive time-out as disciplinary strategy 4. Importance of sharing emotions, empathy, and understanding of the child when using discipline. The last 2 booster sessions reviewed the most important aspects of the above with regards to the needs and experiences of each dyad during the intervention. A summary booklet of all topics covered was issued to each mother at the last session. In each session interveners ensure mothers are treated as experts on their own children, and any examples of posi-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
effect was measured by repeated multivariate analysis of variance for harsh discipline and parenting stress, with 'experimental condition as a between-subjects factor and time as a within-subject factor' (p609). Pretest parenting stress was then tested as a potential moderating factor for intervention effect on harsh discipline. **Country:** Not UK. Portugal. **Source of funding:** Government. Netherlands Organisation for Health Research and Development.	tive mother-child interactions or effective parenting strategy are positively reinforced. Location/place of delivery Own homes of families. Describe comparison intervention Control group mothers received 6 phone calls, made at the same intervals as VIPP-SD home visits, about topics of child development- language, play, sleep, feeding, positive relationships and in the last call, an overview of them all. Mothers were asked questions and encouraged to talk about their child's development but there was no feedback or advice provided by the researcher. Those mothers who asked for explicit advice or help were signposted to their GP and/or health agency. **Outcomes measured** Incidence of abuse and neglect Harsh discipline, measured using standardised procedures adapted from Verschueren et al. (2006). Quality of parenting and parent-child relationships Parenting stress measured using the Daily Hassles Questionnaire (Kanner et al. 1981). **Follow-up**		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	One month after last intervention home visit or phone call. **Costs** No.		

16. Robling M, Bekkers M-J, Bell K et al. (2015) Effectiveness of a nurse-led intensive home-visitation programme for first-time teenage mothers (Building Blocks): a pragmatic randomised controlled trial. Lancet: 1–10

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To investigate the effectiveness of a nurse-led home visiting programme for first-time mothers aged 19 years or younger in a UK context 24 months after birth. The primary aims of the intervention were to reduce tobacco use by the mother, proportion of subsequent pregnancies within the 24 months, birthweight of the baby, and number of emergency department attendances and hospital admissions. We have also extracted data collected	**Participants** Children and young people - Children up to 2 years old. Caregivers and families - First-time mothers aged under 19 recruited at less than 25 weeks gestation. Women who had had previous pregnancies but which ended due to miscarriage, stillbirth or termination were eligible. **Sample characteristics** • Age - Women were recruited at aged 19 and younger. Mean age for both intervention and control groups was 17.9. • Sex – Female. • Ethnicity - Intervention group (n=808) 88% (n=711) White 6% (n=47) Mixed 2% (n=16) Asian 4% (n=31) Black <1% (n=3) Other Control group (n=810) 88% (n=714) White 5% (n=42) Mixed 1% (n=11)	**Effect sizes** Incidence of abuse and neglect Safeguarding procedures 24 months after birth (n=945) a higher proportion of children in the FNP group had a safeguarding event noted in their GP record (n=64/469 (13.6%) v 38/476 (8.0%) AOR 1.85, (CI 95% 1.02 to 2.85) p=0.005. Referral to social services 24 months after birth a higher percentage of participants in the intervention group reported that their child had ever been referred to social services (n=119/580, 20.5% vs. n=91/541, 16.8%): AOR 1.27 (CI 0.93 to 1.73) p=0.13 There was no statistical difference between the groups for IPV and Child safety outcomes. IPV - AOR of 1.17, (CI 95% 0.84 to 1.63, p=0.37) Risk of abuse and neglect Parental role strain - a very small effect was seen, with lower mean score in the FNP group: adjusted difference in means, = -0.16, (CI 95% -0.35 to 0.03,) p=0.11 Quality of parenting and parent-child relationships	**Overall assessment of internal validity** + Bias of using only risk factor of age of mother. Different groups of sample used for different outcomes. **Overall assessment of external validity** + Due to mismatch in primary outcomes. **Overall validity score** + This study has been rated as moderate due to the rigorous design and detailed reporting of data. However, we considered rating this

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
by the study, which relates to outcomes of interest to this review. **Methodology:** RCT inc cluster. Non-blinded, parallel group RCT, with women randomly assigned to either the FNP intervention or usual care. Women from 18 different English local authority-primary care-secondary care partnership areas were identified and approached via local maternity services and were in the main recruited at their home by locally-based researchers. Randomisation was stratified by site and minimised by gestation (less than 16 weeks vs. 16 weeks and over), smoking (yes or no) and preferred language for data collection (English vs. non-English). It was also 'weighted towards minimising the imbalance in trial	Asian 5% (n=50) Black <1% (n=3) Other • Religion/belief - Not reported • Disability - Not reported • Long term health condition - Not reported • Sexual orientation - Not reported • Socioeconomic position – Not in education or training (NEET) Intervention group (n=695 as only applicable to those older than 16 years at end of last academic year) 48% (333/695) Yes 52% (362/695) No Control group (n=685 for reasons as above) 48% (330/685) Yes 52% (355/685) No Paid job Intervention group (n=808) 21% (n=174) Yes 78% (n=634) No Control group (n=810) 20% (n=164) Yes 80% (n=646) No In receipt of government welfare payments Intervention group (n =808) 37% (n=301) Yes 63% (n=507) No Control group (n=808 (n=2 no data available)) 35% (n=283) Yes 65% (n=525) No Ever been homeless Intervention group 18% (n=144) Yes 82% (n=664) No Control group 21% (n=170) Yes 79% (n=640) No Highest parental qualification Intervention group (n=805) 13% (n=108) Up to post graduate 21% (n=172) Up to A level 10% (n=79) Overseas or other qualification 16% (n=130) None of these 39% (n=316) Don't know Control group 13% (n=108) Up to post graduate	There was no significant differences between the groups for maternal-child interaction outcomes (intervention effect is calculated from adjusted difference in means). Maternal sensitivity scores - = -0.07 (CI 95% -0.41 to 0.27), p=0.67 Maternal intrusiveness scores =0.12 (CI 95% -0.19 to 0.43), p=0.44 Child responsiveness score = -0.26 (CI 95% -0.77 to 0.25) p= 0.31 Child positive effect score = -0.23, CI 95% -0.59 to 0.13), p=0.21 Child negative effect score=0.09 (CI 95% -0.12 to 0.30) p=0.40. Children and young people's health and wellbeing outcomes Attendance and admission to emergency department/hospital for injuries and ingestions 587 (81%) out of 725 assessed children whose mothers were in the FNP group and 577 (77%) of 753 children from the comparison group attended A&E or were admitted to hospital at least once before their second birthday (AOR 1.32, CI 97.5% 0.99-1.76). Full scores: Primary care consultation at 24 months AOR =0.87, (CI 95% 0.58 to 1.33), p=0.53 A greater proportion of children in the intervention group than the control group attended A&E for injury or ingestion at 6 months and 24 months A&E attendance at 6 months - AOR=1.52 (CI 95% 0.86 to 2.70) p=0.15 A&E attendance at 24 months - AOR=1.16 (CI 95% 0.92 to 1.46) p=0.20. However a smaller proportion in the intervention group were admitted to hospital for injury or ingestion at 6 and 24 months. Hospital admission at 6 months - AOR=0.79 (CI 95% 0.39 to 1.60) p=0.51 Hospital admission at 24 months - AOR=0.72 (CI 95% 0.46 to 1.12) p=0.15 Looking at both attendance and admission there was no statistical difference between the two groups. Child safety AOR=1.26, (CI 95% 0.97 to	as poor for two reasons: 1) Possible risk of bias arising from high levels of missing data. In line with the Cochrane guidelines we have considered whether the missing data presents a risk of bias in terms of 'the amount and distribution across intervention groups, the reasons for outcomes being missing, the likely difference in outcome between participants with and without data, what study authors have done to address the problem in their reported analyses, and the clinical context' (http://handbook.cochrane.org/). For the safeguarding data provided by GPs our view is that the risk of bias resulting from attrition is low because rates of returned information from GPs is similar for both intervention and control. For self-report data we

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
groups with probability 0.8' (p3). Allocation was concealed using a remote computer-based system, accessed via telephone or internet by the recruiting researcher. Routine data (on birthweight, emergency department attendances and admissions and second pregnancies, as well as for some secondary outcomes) was collected by field-based researchers from maternity units; downloaded by a trial statistician from the Health and Social Care Information Centre (HSCIC); by field-based researchers or practice staff from primary care centres; from the Abortions Statistics Manager at the Department of Health for abortion statistics, from COVER (Coverage Of Vaccinated Evaluated Rapidly) contacts directly from	21% (n=172) Up to A level 10% (n=79) Overseas or other qualification 16% (n=130) None of these 39% (n=316) Don't know • Type of abuse – Not reported. • Looked after or adopted status – Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not applicable. **Sample size** Comparison numbers - n=810 received usual care Intervention numbers - n=808 were randomised to receive the FNP 'Building Blocks' programme Sample size - 1645 randomised n=823 intervention group n=822 usual care. **Intervention** Intervention category – Family Nurse Partnership. Describe intervention - 'Building Blocks' programme was an adapted version of the Nurse Family Partnership (NFP) which is a programme of home visits made to younger first time mothers by specially trained nurses from early pregnancy until the child is two years old, designed to improve pregnancy outcomes, child health and development, a reduction	1.62) p=0.08. Cognitive development at 12, 18 and 24 months, There was no statistical difference seen between groups at 12 and 18 months with regards to developmental concerns; at 24 months though fewer mothers in the intervention group reported concerns than in the control (n=46/569 (8.1%) vs. n=66/522 (12.65%)) Cognitive development concerns at 12 months - AOR 0.91 (CI 95% 0.59 to 1.40) p=0.66 Cognitive development concerns at 18 months - AOR 0.59 (CI 95% 0.32 to 1.11) p=0.10 Cognitive development concerns at 24 months - AOR 0.61 (CI 95% 0.40 to 0.90) p=0.0.13 Language development at 12, 18 and 24 months While more mothers in the FNP arm reported language concerns about their children at 12 months, at 18 months concerns were at similar levels, while at 24 months language development was scored higher among children in the intervention group v the control group. Language development concerns at 12 months - AOR 0.50 (CI 95% 0.35 to 0.72) p=<0.001 Language development concerns at 18 months -AOR 0.66 (CI 95% 0.48 to 0.90) p=0.009 Early Language Milestone score at 24 months -Adjusted difference in means = 4.49 (CI 95% 0.52 to 8.45) p=0.027 Caregiver/parent health and wellbeing outcomes General self-efficacy A small difference was seen in self-efficacy scores, which were higher in the intervention group was seen: adjusted difference in means=0.44 (CI 95% 0.10 to 0.78) p=0.011 There were no other statistical differences seen for parent outcomes: Psychological distress AOR=-0.39 (CI 95% -1.19 to 0.40) p=0.33 Depressive symptoms AOR = 0.80 (CI 95% 0.60 to 1.05) p=0.11 Postnatally depressed (score>13) AOR = 1.03 (CI 95% 0.69 to	suggest that there may be some risk of bias due to higher levels of failure to complete assessments at 24 months amongst the control group (21.9% incomplete) compared to the intervention group (14.3% incomplete). The reviewing team have calculated this difference as being significant (chi-square=13.6, p<0.01). 2) The validity of the study outcome measures in terms of our principal outcome of interest - incidence of abuse and neglect. Safeguarding concerns are measured in the study via GP records and self-report. We are concerned that both of these methods are likely to under-report safeguarding issues, and consider that a better outcome measure would have been to use children's social care data. However, the problem of under-

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
primary healthcare authorities. Data about emergency attendances and admissions and second pregnancies was also collected from mothers. Information about tobacco use was collected via self-report and urine sample- cotinine levels were measured before allocation and at 34-36 weeks gestation. At both times, women were classified as non-smokers if they reported not smoking in the 3 days prior to the interview and had a urinary cotinine concentration of less than 100 ng/ml. if only baseline cotinine levels were collected, baseline self-reports and cotinine concentrations were compared to assess accuracy of self-reporting- accurate and over-reporting mothers were accepted as non-smokers. Local researchers administered a face-to-	in child maltreatment and increase maternal self-sufficiency. Delivered by Family nurses (Band 7) were health visitors, nurses and midwives trained by the FNP central team using a core FNP learning programme (including residential, team-based and specialist master class modules). They were all educated to degree level (or equivalent) all held a nursing or midwifery qualification, were registered with the Nursing Midwifery Council of the UK, masters education desirable. Supervisors (Band 8a) were expected to meet the same criteria but with essential master's education. As part of the adaptation to be delivered in England, specialist supervision including regular psychological support was arranged for Family nurses, safeguarding supervision and systems were implemented and FNP was incorporated into local governance arrangements. Delivered to First-time mothers aged 19 or younger. Duration, frequency, intensity, etc. Series of a maximum of 64 visits which should commence early in the second trimester and decrease in fre-	1.52) p=0.90 Unplanned antenatal hospital admissions IRR =1.0 (CI 95% 0.10 to 0.78) p=0.49 A&E attendances and admissions IRR=1.26 (CI 95% 0.10 to 0.78) p=0.07 Foster care for the mother Data suppressed on low cell counts.	reporting should be equally present in both intervention and control groups. Based on these considerations we decided to retain a rating of 'moderate'.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
face structured computer-assisted personal interview before allocation and then 24 months after birth. This interview gathered self-reported data on secondary outcomes These researchers were not blind to allocation but were independent of the service delivered. Computer-assisted telephone interviews were conducted by office-based researchers (who were blind to allocation) at late pregnancy (34–36 weeks gestation), 6, 12 and 18 months after birth. Mothers and family nurses were asked to report contact with health services and their clients respectively. Mothers were able to withdraw at any time, and miscarriage, death or adoption of the child lead to mandatory withdrawal. Those that wanted to discontinue the intervention were	quency over time until the child's second birthday. Number of visits may vary according to individual need, engagement, and gestational age at enrolment, but there are minimum targets to support particular desired outcomes. Key components and objectives of intervention NFP is based on theories of human ecology, self-efficacy and human attachment. (p5 of appendix). Core content for visits: personal and environmental health; life course development; maternal role; family and friends; access to health and social services. (Each area is allocated a specific amount of time by the programme). Maternal self-efficacy is promoted throughout the programme as is sensitive and competent care giving through education and modelling activities using a strengths-based approach, in order to reduce maltreatment. In order to adapt to the UK, the language of the programme was changed to UK English. A core element, introduced by the University of Colorado to the FNP model, is now motivational interviewing, and nurses are now trained to use a 'mainly guiding communication style' with the young mothers. (p5 appendix).		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
still asked to consider giving follow-up data. At 24 months after birth all participants were invited to provide follow-up data, even those who did not provide data at other preceding data collection waves. Adverse events (notified to the trial office) were reviewed by a senior clinical researcher for relatedness and severity. **Country:** UK.18 partnerships between local authorities and primary and secondary care organisations in England. **Source of funding:** Government. Department of Health who stipulated some policy-related primary outcomes for the study but had no input on the study design.	Location/place of delivery Home Describe comparison intervention Usual care- the Healthy Child Programme (HCP) universal offer of clinical and public health for children and families, involving home visits from health visitors. **Outcomes measured** Incidence of abuse and neglect Safeguarding, defined as any record in GP notes indicating the initiation, progression or closure of a safeguarding process. Referrals to social services Intimate partner violence (measured using the Composite Abuse Scale (Hegarty 2007) at 24 months, in face to face interviews when participant was alone, scoring up to 145, a score of 0 indicating no abuse); Risk of abuse and neglect Parental role strain, (measured at 6, 12, 18 and 24 months using Millennium Cohort Study variables, IoE 2003 and 2004, 6-item scale with scores 6-24, low score indicating lower parental strain); Quality of parenting and parent-child relationships Maternal-child interaction (measured using scale adapted from Fish &		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Stifter 1995 analysis of mother child-interaction and attachment). Children and young people's health and wellbeing outcomes Attendance and admission to emergency department/hospital for injuries and ingestions (measured using hospital episode statistics or GP provided data) Child safety (measured using scale adapted from Health of California's Adults, Adolescents and Children, findings from CHIS 2003 and CHIS 2001,Holtby et al. 2006); Cognitive, language a development at 12, 18 and 24 months, measured using Schedule of Growing Skills II (SOGS II) Bellman et al. 1996 at 12 and 18 months, and the Early Language Milestone scale, (Coplan et al 1982) at 24 months. Caregiver/parent health and wellbeing outcomes Psychological distress (measured using 10-item Kessler scale, Kessler et al 2002, scores 10 to 50, low score indicating low level of distress) Depressive symptoms (measured using Whooley scale, Whooley et al. 1997, where positive responses to questions indicate presence of symptoms) General self-efficacy (measured using the 10-item General Self-efficacy Scale, scores ranging 10 to 40, where higher score indicates higher efficacy) Unplanned hospital admissions, A&E		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	attendances and admissions and foster care for the mother (using data from primary care and statutory services). **Follow-up** Data was collected at 6 months, 12 months, 18 and 24 months after birth. **Costs** 'Full-sample base case analysis (n=782 FCP vs. n=786 in usual care group) with use of multiple imputation showed an incremental cost for FNP of $1993 per participant. Although individual types of resource use were similar across trial group, intervention delivery costs, (FNP calls and visits) accounted for incremental cost of FNP. Sensitivity analysis that included complete cases only (n=217 in FNP group v 186 in usual care) suggested the incremental cost of the FNP was £4670 (95% CI 3322-6017)' (p6).		

17. Sanders MR, Pidgeon AM, Gravestock F et al. (2004) Does parental attributional retraining and anger management enhance the effects of the Triple P-Positive Parenting Program with parents at risk of child maltreatment? Behavior Therapy 35: 513–35

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The study aimed to compare the	**Participants:** • Children and young people.	**Effect sizes - Risk of abuse and neglect:** (Exact val-	**Overall assessment of internal validity:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
effectiveness of '... an enhanced group behavioral family intervention (EBFI) for parents at risk of child maltreatment ...' which addressed parental negative attributions and anger management to a '... standard-care group parent training intervention ...' (p516). **Methodology:** RCT. **Country:** Not UK - Australia.	Children between the ages of 2 and 7 with a parent experiencing anger management problems in relation to their child. • Caregivers and families. Parents experiencing anger management problems in relation to their child. **Sample characteristics:** • Age - Mean age of child in months (*SD*): SBFI 53.71 (19.32); EBFI 52.84 (17.85). Mean age of mother in years (*SD*): SBFI 33.29 (5.35); EBFI 33.68 (5.58). Mean age of father in years (*SD*): SBFI 35.32 (6.34); EBFI 36.45 (7.14). Mean age of participating parent in years (*SD*): SBFI 33.33 (5.37); EBFI 34.18 (6.34). • Sex - Female parent participant: SBFI 92%; EBFI 94% Female target child: SBFI 52%; EBFI 48%. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Mean number of children (SD): SBFI 1.92 (0.87); EBFI 2.38 (1.31). Mean number of years together as a couple (*SD*): SBFI 7.78 (3.93); EBFI 9.38 (4.91). Married: SBFI 73%, EBFI	ues of statistical tests are recorded here only for significant time x condition interactions.) Risk of maltreatment was measured using five scales which all relied on parental report:- The Blame and Intentionality subscale of the Parent's Attributions for Child's Behavior (Pidgeon & Sanders 2002): Ambiguous situations: Significant interaction of time and condition in favour of EFBI (F=6.42, p>0.01) Intentional situations: Significant interaction of time and condition in favour of EFBI (F=7.72, p>0.001). The State-Trait Anger Expression Inventory (Spielberger 1996). No significant difference between conditions, although main effect of time. The Parental Anger Inventory: No significant difference between conditions, although main effect of time. The abuse subscale of the Child Abuse Potential Inventory (Milner 1986): Significant interaction of condition with time in favour of EFBI (F=4.82, p<0.01). The Parent Opinion Questionnaire (Azar and Rohrbeck 1986). Significant interaction of condition with time in favour of EFBI (F=5.06, p<0.01). **Effect sizes - Quality of parenting and parent-child relationships:** Parenting was measured using 2 scales, both of which rely on parental self-report. Neither scale showed significant difference between conditions, or a time x condition interaction, although there was a main effect of time. **Effect sizes - Children and young people's health and wellbeing outcomes:** Child behaviour measures - no significant differences between conditions, or time x condition interactions. Although there was a main effect of time. Using three criteria to assess clinical sig-	**Overall assessment of external validity:** + **Overall validity score:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	66% Incomplete secondary education: SBFI 44%; EBFI 60% Primary occupation of 'home duties': SBFI 58%; EBFI 55% Annual family income of less than $25,000 (AUD): SBFI 25%; EBFI 31% Financial difficulties in family: SBFI 34%; EBFI 25% Participants who currently use 'illicit' drugs: SBFI 6%; 6% Participant who currently abuses alcohol (>40g/day): EBFI 6%; 0%. • Type of abuse - Not reported but some families had already been referred to authorities for allegations (did not need to be substantiated) of abuse or neglect at baseline. Contact with statutory authority for suspected abuse or neglect: SBFI 4%; EBFI 6%. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - Standard behavioral family intervention programme (SBFI): Pre-intervention n=48; post-intervention not clear; 6 month assessment n=39. • Intervention numbers - Enhanced behavioral family intervention programme (EBFI): Pre-intervention	nificance of change in children's behaviour also revealed no significant differences between EFBI and SFBI at either immediately post intervention or 6 month follow-up. **Effect sizes - Caregiver/parent health and wellbeing outcomes:** Parental adjustment - no significant differences between conditions, or time x condition interactions, although main effect of time. **Effect sizes - Satisfaction with services:** Not able to find results of client satisfaction measure. **Narrative findings – Effectiveness:** This study suggests that parents participating in both variants of the Triple P intervention showed significant improvements in a wide range of measures of family functioning, with EFBI showing significantly better performance in a number of areas. For example, EFBI participants showed a significantly greater reduction in two measures of child abuse risk: child abuse potential (CAPI scores) and unrealistic expectations (POQ) scores. EFBI participants also showed a reduction in dysfunctional attributions than the comparison group. There was significant improvement in both conditions, but no significant difference between conditions for: - Parental angry temperament - Parental anger inventory - Parenting scale - Parenting sense of competency - Child behaviour measures, including when assessed for clinically significant change - Parental adjustment.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	n=50; post-intervention not clear; 6 month assessment n=35. **Intervention category:** Parenting programmes. **Intervention:** The Enhanced behavioral family intervention programme (EBFI) usually took around 12 weeks to complete and was essentially the same as the SBFI (see information below on comparison intervention) with the addition of – 4 group sessions (2 hours in duration) of parent training. Four sessions (2 hours in duration) targeted at 'additional risk factors' associated with abuse and neglect. Parents were given a workbook which covered the principles taught in these sessions. The sessions aimed to teach parents skills which would help them to challenge their own beliefs regarding their own behaviour and that of their child, and to address negative parenting strategies associated with these beliefs. Cognitive and physical planning strategies for anger management were also introduced; although it should be noted that both treatment conditions addressed the issue of planning ahead for high risk situations and encouraged parents to develop coping plans accordingly.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Fourteen practitioners (12 female and 2 male) were trained and supervised in the delivery of the interventions (1 clinical psychologist; 8 psychologists completing postgraduate training in psychology; 2 psychologists; 2 social workers; and 1 teacher). **Comparison intervention:** The Standard behavioral family intervention programme (SBFI) usually took around eight weeks to complete. Families received 4 group sessions (2 hours in duration) of parent training; four individual telephone consultations after completion of group sessions (15–30 minutes in duration); and a copy of 'Every Parent's Group Workbook (Markie-Dadds et al. 1997) covering the main principles of the programme and associated exercises. The programme was composed of 17 child management strategies focusing on children's competence and development and management of misbehaviour. Parents were also taught '… a planned activities routine to enhance the generalization and maintenance of parenting skills' (p523). This planning for high risk situations e.g. when the child is bored and to plan activities appropriate to the child's age. Parents also learnt how to monitor and set		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	goals for behaviour change and to improve their observational skills in relation to their own and their child's behaviour. The authors describe the training as active and note the use of techniques such as rehearsal, and goal setting. **Outcomes measured:** • Risk of abuse and neglect - Risk of maltreatment was measured using five scales which all relied on parental report:- The Blame and Intentionality subscale of the Parent's Attributions for Child's Behavior (Pidgeon & Sanders 2002). The State-Trait Anger Expression Inventory (Spielberger 1996). The Parental Anger Inventory (Hansen & Sedlar 1998) yielding a problem and intensity score. The abuse subscale of the Child Abuse Potential Inventory (Milner 1986). The Parent Opinion Questionnaire (Azar and Rohrbeck 1986). • Quality of parenting and parent-child relationships - Parenting was measured using two scales, both of which rely on parental self-report:- The Parenting Scale (Arnold et al. 1993). Parent Sense of Competence (Gibaud-Wallston and Wandersman, 1978). The study also measured 'parenting contexts for child behavior		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	problems' via the Home and Community Problem Checklist (Sanders and Dadds 1993). • Children and young people's health and wellbeing outcomes - Child behaviour was measured using two scales, both of which relied on parental reports:- The Eyberg Child Behavior Inventory (Eyberg and Pincus 1999). The Parent Daily Report Checklist (Chamberlain and Reid 1987). • Caregiver/parent health and wellbeing outcomes - Parental adjustment was measured using two scales:- The Depression-Anxiety-Stress Scales (Lovibond and Lovibond 1995). The Parent Problem Checklist (Dadds and Powell 1991). • Satisfaction with services - Self-reported - measured using the Client Satisfaction Questionnaire (Sanders et al., 2000). **Follow-up:** Assessments were conducted at pre and post treatment and at 6 months.		

18. Scudder AT, McNeil CB, Chengappa K et al. (2014) Evaluation of an existing parenting class within a women's state correctional facility and a parenting class modeled from parent–child interaction therapy. Children and Youth Services Review 47: 238–47

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To investigate '... whether a PCIT-based parenting class provides additional benefit when compared to the existing parenting classes in a women's correctional facility in enhancing demonstrated parenting skills as well as parent-reported knowledge of child development, parenting stress, child abuse potential, and treatment acceptability'. **Methodology:** RCT. **Country:** Not UK - USA.	**Participants:** Caregivers and families. Mothers incarcerated in a state correctional facility. **Sample characteristics:** • Age - Mean age of 30.83 years (SD=5.88) - mean ages of intervention and comparison groups were not reported. Average age of all participants' children (i.e., not just target child) was 6.76 years (SD=4.37). • Sex - Sample was 100% female. • Ethnicity - White, non-Hispanic=93% African American=3.7% Biracial=3.7% • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Below poverty line = 72.9% (total sample). Single or never married = 59.5% (intervention), 37.5% (comparison). Divorced or separated = 40.5% intervention, 32.5% (comparison). Had not graduated from high school = 71.5% (intervention), 55% (comparison). Unemployed or in receipt of disability benefits at time of entry into prison = 64.3% (intervention),	**Effect sizes - Risk of abuse and neglect:** Child Abuse Potential - No significant difference between conditions. Parenting Stress Index - total parent stress: F=0.179, p=0.67, d=0.08. Adult–Adolescent Parenting Inventory Inappropriate expectations subscale - significant difference between conditions (F=7.54, p=0.05, d=.51). Examination of mean scores suggests this is in favour of the control rather than intervention group (PCIT pre mean = 20.63, post = 20.95; TAU pre mean = 20.55, post = 22.74). Adult–Adolescent Parenting Inventory empathic regard subscale - no significant difference between conditions. Adult–Adolescent Parenting Inventory discipline subscale - no significant difference between conditions. **Effect sizes - Quality of parenting and parent-child relationships:** Interaction quality as assessed by Dyadic Parent–Child Interaction Coding System-III - Dyadic Parent–Child Interaction Coding System-III -Positive attention 'do' skills: F=38.96, p=0.00, d=1.67. Dyadic Parent–Child Interaction Coding System-III - Negative attention 'don't' skills: F=17.02, p=0.00, d=0.83. Dyadic Parent–Child Interaction Coding System-III - Effective command sequences: F=3.92, p=0.05, d=0.54. Dyadic Parent–Child Interaction Coding System-III - Compliance-contingent praise: F=14.70,	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity score:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	47.5 (comparison). Household income below $15,000: 26.2% (intervention), 25.0% (comparison). • Type of abuse - No clear focus but the authors state in the limitations section that a history of harsh parenting or physical abuse of the child was not required for participation. • Looked after or adopted status - Both intervention (12%) and comparison (4.3%) groups included mothers who had a child who was now living with a non-family member appointed by a court. 12% (intervention). • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - Baseline n=40, post-treatment n=30. • Intervention numbers - Baseline n=42, post-treatment n=39. • Sample size - n=82. **Intervention category:** Parenting programmes. **Intervention:** Parent Child Interaction Therapy has several core features: 'a) The parent and child are actively involved together in treatment sessions b) Interactions are coded to assess	p=.00, d=1.02. Adult–Adolescent Parenting Inventory - Inappropriate expectations: F= 7.54, p= 0.05, d= 0.51. Adult–Adolescent Parenting Inventory - Empathetic regard: F=.311, p=0.58, d=0.01. Adult–Adolescent Parenting Inventory - Discipline: F=2.33, p=0.13, d=0.13. **Effect sizes - Satisfaction with services:** Therapy Attitude Inventory - Treatment satisfaction: F=2.07, p=0.04, d=0.50. **Narrative findings – Effectiveness:** There were no significant between group differences in scores on either the Child Abuse Potential Inventory or the Parenting Stress Index scales. The authors highlight that total parenting stress amongst mothers in this trial was at a lower level than previously reported in other studies evaluating Parent Child Interaction Therapy. They suggest that lower total scores on the Parenting Stress Index may be due to limited contact with children during imprisonment and a smaller intervention dose than that received in earlier studies. Significant between group differences were found (in favour of the intervention) on all 4 scales of the Dyadic Parent–Child Interaction Coding System which measures the interactions between parents and children. Significant between group differences were found (in favour of the control group) on the 'inappropriate expectations' subscale of the Adult–Adolescent Parenting Inventory. The authors suggest that this may be due to the fact that '... child development is not a component of parent-training programs associated with program effect	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	progress and determine treatment planning c) Parents are coached to assist in reaching a level of mastery of both play-therapy and discipline skills d) Traditional play-therapy skills are taught to enhance the quality of the parent-child relationships e) Parental problem-solving skills and use of behavioural principles are coached to develop direct strategies for management of problem behaviours f) Changes are made on empirical evidence' (p240). Parenting classes (12–15 participants received no more than 7 classes) based on Parent Child Interaction Therapy delivered in a group format by an instructor and an undergraduate assistant. Classes lasted for 90 minutes (total of 10.5 hours) and included discussion of PCIT protocol topics, coaching via parenting role-play, peer review of other participants PCIT skills and assignment of tasks to be practiced outside of the class (with another participant or their children where possible). It may be important to note that the intervention was modified to enable delivery in a correctional environment (e.g. in conventional PCIT	sizes for parent behavior ...' (p245). No significant differences were found between groups using the 'empathetic regard' and 'discipline' subscales of the Adult–Adolescent Parenting Inventory. There were significant between group differences (in favour of the intervention) on measures of treatment satisfaction.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	progression usually depends on 'mastery' of certain skills. In this program this was not required for progression or graduation.) Participants also had access to toys and equipment available for family visits. **Comparison intervention:** Assignment to the existing parenting programme available at the facility. Weekly meetings of (total of 10.5 hours) led by an instructor and an assistant (an inmate who had already completed the programme). The programme was adapted from the Partnerships in Parenting manual and usually included discussions and presentations by the instructors on specific topics such as discipline. Role-play was used in some classes and participants were given 'homework' assignments. **Outcomes measured:** • Risk of abuse and neglect - Child Abuse Potential Inventory (CAP, Milner 1986) Adult–Adolescent Parenting Inventory, 2nd edn (AAPI-II, Bavolek & Keene 1999). • Quality of parenting and parent-child relationships - Interaction quality as assessed by Dyadic Parent–Child Interaction Coding System-III		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	(DPICS-III, Eyberg et al. 2005) Parenting Stress Index, Third Edition (PSI, Abidin 1995). • Satisfaction with services - Therapy Attitude Inventory (TAI, Eyberg & Johnson 1974). **Follow-up:** Outcomes were measured at pre and post treatment.		

19. Silovsky JF, Bard D, Chaffin M et al. (2011) Prevention of child maltreatment in high-risk rural families: A randomized clinical trial with child welfare outcomes. Children and Youth Services Review 33: 1435–44

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: 'To conduct a randomised clinical trial of SafeCare augmented for rural high-risk population (SC+) compared to standard home-based mental health services (SAU) to examine reductions in future child maltreatment reports, as well as risk factors and factors proximal to child maltreatment' (p1435). **Methodology:** RCT. **Country:** Not UK –	**Participants:** • Caregivers and families - Participants were caregivers aged at least 16, with at least 1 child aged 5 years or younger and at least 1 of the following risk factors: parental substance misuse, mental health issues or intimate partner violence. **Sample characteristics:** • Age - Mean age of intervention group at baseline = 25.9 years; mean age of control group at baseline = 27.7 years. • Sex - Experimental group = 100% female; control group = 98% female. Ethnicity - Ethnicity of participants as follows: Intervention group: White	**Effect sizes - Incidence of abuse and neglect:** Ten (20.8%) of SC+ and 18 (31.5%) of SAU participants had a future referral to child welfare. It is not reported whether this is a statistically significant difference. No significant differences found between intervention and control group on any measures of incidence of abuse, except for reports due to domestic violence (log rank test: chi square = 6.91, $p<0.01$). Effect sizes not reported for any variables. **Effect sizes - Risk of abuse and neglect:** Parent-Child Conflict Tactics Scale - Only one subscale fit the generalised linear mixed model: nonviolent discipline. This showed a significant improvement in favour of SC+ between pre- and post-service, but this was not sustained at follow-up. All other scales showed no significant impact.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity score:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
USA. **Source of funding:** Government.	68%; Black or African American 15%; Hispanic or Latino 2%; American Indian or Alaska Native 15%; Asian 0% Control group: White 74%; Black or African American 14%; Hispanic or Latino 4%; American Indian or Alaska Native 7%; Asian 1% • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - In employment: Intervention group: Yes 54%; No 46% Control group: Yes 56%; No 44%. • Type of abuse - Not reported. Caregivers with substantiated report of child sexual abuse excluded from study. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - Baseline: 57 Received SAU: 19 Wave 2 post-service: 12 lost to follow-up Wave 3 follow-up: 12 lost to follow-up. See p. 1437 for diagram. • Intervention number - Baseline: 48	Effect sizes for the above are not reported. A web address is given for more information, but link is incorrect. **Effect sizes - Satisfaction with services** Significant or near-significant between-group differences shown in favour of intervention group on all measures of service satisfaction and engagement. Effect sizes not given (t test only). Satisfaction: t=-4.0, p<0.05 Cultural competency: t=-1.9, p<0.10 Completed intake: t=4.8, p>0.001 Billable hours: t=5.0, p<0.001. **Narrative findings – effectiveness:** The study reports that parents in the intervention group self-reported (in the Client Satisfaction Survey) greater improvements in parenting behaviours than the control group. However, 'this did not translate into significant group differences in self-reports of discipline strategies, risk and protective factors, or reports to child welfare' (p1443). There were significant or near-significant between-group differences shown in favour of intervention group on all measures of service satisfaction and engagement. The authors note that, given lack of engagement is a key problem for services of this nature, this is a positive outcome. Although the differences in child welfare outcomes were non-significant, the study reports that these were 'promising' and 'in the right direction' (p1443).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Received SC+: 40 Wave 2 post-service: 5 lost to follow-up Wave 3 follow-up: 14 lost to follow-up. See p. 1437 for diagram. • Sample size - Intervention: n = 48 Control: n=57. **Intervention category:** Home visiting - SafeCare. **Intervention:** SafeCare is based on an eco-behavioural model of maltreatment (Lutzker 1984). The intervention is ecological, in that it targets different levels within a concentric ecological model of maltreatment. It is behavioural in relation to emphasising proximal skills and behaviours and 'training parents to criterion in observable skills' (p1436). SC targets parenting behaviours in relation to child health, home safety and cleanliness, and parenting-child bonding. SafeCare augmented (SC+) comprises SC with the addition of Motivational Interviewing (Miller and Rollnick 2004), and training home visitors on identification and response to child maltreatment and risk factors. The service is provided by home visitors who are trained and observed for fidelity to the model by certified monitors. In this intervention, providers were also trained in Motivation Inter-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	viewing by a member of the Motivational Interviewing Network of trainers. There is no information in the study regarding the professional background or other skills of home visitors. In this study, those in the SC+ group received an average of 36 hours of intervention, and those in the SAU group received an average of hours of intervention. It is unclear if this is the intended 'dosage' of either intervention. **Outcomes measured:** • Incidence of abuse and neglect - Measured via child welfare referrals and child removal data. These were measured using 2 'survival' outcomes: - time to first report of any abuse or neglect - time to first report of neglect. Risk of abuse and neglect - Risk and proximal maltreatment factors measured via: 1. Child Abuse Potential Inventory (Milner 1986) 2. Conflict Tactics Scale - Parent-Child Version (CTS-PC) (Straus et al. 1998) 3. Beck Depression Inventory (Beck 1996) 4. Conflict Tactics Scale 2, measuring adult-to-adult conflict (Straus et al. 1996) 5. Family Resources Scale - revised (Dunst & Leet 1987), measuring adequacy of		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	resources in households with children. • Satisfaction with services - Measured via: 1. Client Cultural Competence Inventory (Switzer et al. 1998) 2. Client Satisfaction Survey. • Service outcomes - Monthly Service Utilisation Report, primarily monitoring delivery of SC+ and MI. **Follow-up:** At two time points - 1. Post-services (no earlier than 6 months after pre-service interview); 2. Follow-up (no earlier than six months after ends of services).		

20. Stover C (2015) Fathers for Change for Substance Use and Intimate Partner Violence: Initial Community Pilot. Family Process 54: 600–9

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The assess the feasibility and efficacy of the Fathers for Change programme for men with history of intimate partner violence (IPV) and substance abuse, compared to individual drug counselling. **Methodology:** RCT inc cluster. RCT- 18	**Participants** Children and young people - Children- biological children under the age of 10, who as the oldest child either living with the participating father or seeing them more than monthly, were given permission to participate by their mothers. Caregivers and families - Fathers (and later, their female co-parents) referred after domestic violence and drug charges.	**Effect sizes** Men in the FFC showed significantly less intrusiveness in play sessions with their child than those in the IDC group (mean scores in this group increased for intrusiveness) (F=.875, p=.365, over time F=7.88, p=.01). FFC fathers also showed more consistency in parenting style than those in the IDC group (F=.78, p=.407, over time F=4.24, p=0.08). There was no significant difference seen in self-reported co-parenting experiences and/or behaviours either over time or as a result of the intervention. Incidence of abuse and neglect	**Overall assessment of internal validity** - Mentions not all participants finished treatment (67% intervention vs. 33% control) but not how this affected the analysis or provides break down of attrition rates. Does not comment on effect size, just significance.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
out of 35 men with histories of both IPV and substance abuse, referred by the US courts or Department of Children and Families (DCF) after domestic violence or drug charges, were randomised, after baseline assessment, to receive the Fathers for Change intervention or individual drug counselling as a comparison group. Participants were included if they a) met the DSM-IV criteria for substance abuse (alcohol, cocaine, or marijuana) consumed within 30 days prior to screening; was reported to have inflicted physical violence in an intimate relationship within the 90 days prior to screening (according to court/police/or self-reports); were the biological father of at least one child under the age of 10 with whom they either lived or had more than	**Sample characteristics** • Age - Father's mean age was 30.19 (SD=6.90); target child's mean age = 3.o5 (SD= 2.78). • Sex - n=18 fathers, n=10 co-parenting mothers. • Ethnicity - Ethnicity- 52% African-American, 14% European American, 19% Latino, 10% Multi-ethnic, 5% other. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - 62% fathers employed at least part-time with on average 11.67 years of education. • Type of abuse – IPV and poor parent-child interaction. • Looked after or adopted status – N/A - only looked at biological children living with either father and mother, or mother with several visits from father a month. • Unaccompanied asylum seeking, refugee or trafficked children - Not applicable. **Sample size** Comparison numbers - n=9 Intervention numbers - n=9 Sample size - n=18 out of 35 eligible men who were included after screen-	Both groups reported a significant reduction in aggression, with a particular reduction in aggression by participant seen in the FFC group (F=3.73, p=0.035) and in participants reporting violence from their partner (F=3.67, p=0.037). There was a marginally significant group x time interaction for IPV (F=1.23, p=0.07). Effect size for group x time interaction calculated by review team using F values is d=0.5228. Quality of parenting and parent-child relationships Men in the FFC showed significantly less intrusiveness in play sessions with their child than those in the IDC group (mean scores in this group increased for intrusiveness) (F=0.875, p=0.365, over time F=7.88, p=0.01). FFC fathers also showed more consistency in parenting style than those in the IDC group (F=0.78, p=0.407, over time F=4.24, p=0.08). There was no significant difference seen in self-reported co-parenting experiences and/or behaviours either over time or as a result of the intervention. Reviewing team calculated effect sizes for significant results: Play intrusiveness group x time interaction: d=1.3233 Consistency in parenting style group x time interaction: d=0.9707.	Small sample but pilot study. **Overall assessment of external validity** + US focus, lack of clarity on setting. **Overall validity score** - Internal validity issues prevent higher rating

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
monthly visitation (if more than one child, the oldest was chosen as participant of assessment and treatment). The female co-parents of eligible participants were contacted (via information from the courts of DCF) to inform them, screen them, invite their participation in later sessions (if co-parent in the intervention group) and gain their consent for shared children to participate also. All gave permission for children to participate, 10 agreed to participate themselves. Men were excluded if they reported suicidal or psychotic symptoms; had a history of bipolar or psychotic disorders, if review of police records and interview with female co-parent revealed significant use of coercive control; had a history of severe violence (strangulation;	ing referrals, agreed to participate, attended assessment and gave informed consent and were not excluded after further assessment (due to lack of or severity of violence in the relationship), n=9 in each group. **Intervention** Intervention category – Parenting intervention. Describe intervention - Fathers for Change is a 4 month programme designed to support fathers, with a history of IPV and co-occurring substance abuse, to cease aggression, violence and substance use, to improve their co-parenting, to alter their parenting behaviours in a positive way, through treatment sessions involving them and their child (aged under 10 years), and potentially the co-parenting mother of their child. Delivered by Therapists - intervention and control drug counselling. Research assistants - administering self-reporting questionnaires. Delivered to Fathers, their children and their co-parents. Duration, frequency, intensity, etc.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
use/threats of with a weapon, threats to kill) female co-parents reported a fear of the father or if they stated they did not want their child participating. Assessment at baseline was via a series of questionnaires completed by fathers and their children's mothers (on separate days and receipt of $50) on severity of violence, severity of substance abuse, parenting, psychiatric issues and the father-child relationship. Fathers were observed during a play assessment with the child chosen to participate. Urn randomisation software was used to randomise participants to ensure comparability of each group. During the intervention period, male participants met with research assistants (blinded to their group status) weekly to com-	60 minute session covering 14 topics over 16 weeks/4 months. Key components and objectives of intervention There are three phases - individual sessions with the father, designed to achieve abstinence from aggressions and substance abuse, co-parenting sessions (if the mother has agreed to participate), designed to improve co-parenting communication, and restorative parenting sessions, designed to improve parenting quality and/or the father-child relationship. Co-parenting mothers are invited to meet the therapist separately after initial assessment and just before the co-parenting sessions begin to ensure they feel comfortable and safe, understand the programme and feel able to voice any concerns about the relationship (after which it will be assessed whether it is safe to begin co-parenting phase). They can participate in up to 6 sessions with the father. During the intervention sessions cover 14 topics which draw on attachment, family systems and cognitive behavioural theory. The intervention builds on the central idea that increasing these men's feelings of competence and meaning as a father will give them the motivation to change their behaviour.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
plete further self-assessment of substance abuse (SA), IPV, and parenting behaviours. Therapy sessions were videotaped and coded for treatment fidelity by trained independent coders, using either the Yale Adherence and Competence Scale-II (Nuro et al. 2007) for intervention group or Adherence-Competence scale for Individual Drug Counselling (Barber et al. 1996). Following intervention, participants completed a posttreatment assessment and a 3 month follow up with research assistants blinded to group status of participants. The Addiction Severity Index was used to measure the severity of substance abuse in the month prior to randomisation; the Timeline Follow-back-Spousal Violence and TimeLine Follow-back-Substance Use	Location/place of delivery Not clear. Describe comparison intervention Evidence-based individual drug counselling with the fathers only, incorporating the disease model of addiction and 12-step treatment. **Outcomes measured** Incidence of abuse and neglect Self-reported aggression and/or physical violence between father and partner. Not specified whether this violence also aimed at children. Quality of parenting and parent-child relationships Improved father-child interaction during play (reduced intrusiveness, improved tone of voice, improved levels of activity and interest), consistency of parenting style, co-parenting undermining and conflict - see methodology for measures. **Follow-up** Posttreatment assessment followed up after 3 months with research assistants blinded to group status of participants. **Costs** No.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
(TLFB_SV and TFLFB-SA, Sobell and Sobell, 1995; Fals-Stewart et al. 2003) was used during the weekly assessments to assess violence and SA during treatment; the Revised Conflict Tactics Scale (CTS2; Straus et al. 1996) was used to assess via self-reporting the amount of violence used by the fathers against the mother of the participating child in the past year. The Co-parenting Relationship Scale (CRS; Feinberg 2003) was used to describe the quality of the co-parenting relationship, drawing on its co-parenting conflict and co-parenting undermining scales for this study. Four tasks were selected from the Child Interactive Behaviour Rating (Feldman 1998), designed to evaluate video-recorded interactions between parents and			

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
their children were used by assessment sessions during this study, altering based on the child's age (2 being developmentally below and 2 advanced). Following 15 minutes of free-play, fathers were asked to introduce their children to each task one at a time, putting the task in front of the child and asking them to complete it. These video-recorded interactions were coded by 2 trained (by the founder) coders using 7 relevant scales from the Child Interactive Behaviour Coding system - 3 adult (intrusiveness; hostility; and consistency) 4 dyadic (tension; reciprocation; fluency and constriction). Participants also completed a modified Client Satisfaction Questionnaire 8 (Donovan et al 2002; Larsen et al. 1979) to note their satisfaction.			

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Repeated measures analysis was conducted to understand the effect of the intervention. **Country:** Not UK. USA. **Source of funding:** Government. National Institute on Drug Abuse.			

21. Thomas R and Zimmer-Gembeck MJ (2012) Parent-child interaction therapy: An evidence-based treatment for child maltreatment. Child Maltreatment 17: 253–66

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To investigate the effectiveness of a Standard Parent–Child Interaction Therapy (S/PCIT) treatment protocol with mothers who were at high risk or who had a history of maltreating their children. **Methodology:** RCT. Part of a larger RCT of PCIT where participants were allocated to	**Participants:** Caregivers and families -Female caregivers and their children. Participants were referred from child protection authorities (34.2%), government health services (19.7%), and education and nongovernment social service organisations (18.4%). Parent self-referrals were also accepted (27.6%). Families referred from child protection were classified as having engaged in child maltreatment. **Sample characteristics:** • Age - Caregivers mean age 33.9 years (SD 7.31), children's mean	**Effect sizes - Risk of abuse and neglect:** Non-significant difference between groups in parent child abuse potential scores from baseline to 12 weeks after intervention. Standard PCIT: Baseline - 153.9, follow-up 137.1 Waiting list: Baseline - 155.1 follow-up 149.1 (ES-Cohen's d = -0.01, p=0.315). **Effect sizes - Quality of parenting and parent-child relationships:** Changes in parent observed behaviours - quality of parents' verbalisations when interacting with their children. • Use of praise at baseline and at 12 weeks after intervention (% of talk) S/PCIT: Baseline 3.6; Follow-up 12.4 Waiting list: Baseline 3.7; Follow-up 4.3 (ES- Cohen's d=1.40; p=0.00).	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity score:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
time-variable PCIT (TV/PCIT), standard PCIT (S/PCIT) or wait-list. (Thomas and Zimmer-Gembeck 2011). Data collection: pre- and post- assessment self-report questionnaires, also videotaped pre-assessment. **Country:** Australia. **Source of funding:** Government - The Future Directions Prevention and Early Intervention Trials, Queensland Department of Child Safety, Australia.	age 4.57 years (SD 1.3). • Sex - Caregivers were all female, children = 70.4% male. • Ethnicity - 74% of caregivers were born in Australia, 1.4% of Aboriginal or Torres Strait Islander descent. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Most mothers had completed some high school (81%) and 16.5% had some tertiary education. • Type of abuse - Different subtypes of maltreatment. Sexual abuse cases were excluded as PCIT has been contraindicated for children with a history of sexual abuse. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - Waiting list n=91 female carers (more families to waiting list group than S/PCIT group due to continuous recruitment from 2002–9). • Intervention numbers - S/PCIT,	• Use of Description/reflection at baseline and at 12 weeks after intervention (% of talk) S/PCIT: Baseline 43.8; Follow-up 61.5 Waiting list: Baseline 45.1; Follow-up 46.8 (ES- Cohen's d=1.28; p=0.00). • Use of questions score at baseline and at 12 weeks after intervention (% of talk) S/PCIT: Baseline 37.3; Follow-up 16.7 Waiting list: Baseline 36.9; Follow-up 35.7 (ES- Cohen's d=-1.50; p=0.00). • Use of commands at baseline and at 12 weeks after intervention (% of talk) S/PCIT: Baseline 13.4; Follow-up 7.9 Waiting list: Baseline 12.8; Follow-up 10.8 (ES- Cohen's d=-0.39; p=0.18). • Use of negative talk at baseline and at 12 weeks after intervention S/PCIT: Baseline 1.7; Follow-up 0.8 Waiting list: Baseline 1.3; Follow-up 1.9 (ES- Cohen's d=-0.61; p=0.002). • Parental sensitivity scores at baseline and at 12 weeks after intervention S/PCIT: Baseline 5.6 Follow-up 6.3 Waiting list: Baseline 5.3 Follow-up 5.4 (ES- Cohen's d=-0.47; p=0.008). **Effect sizes - Children and young people's health and wellbeing outcomes:** Parents' report on child externalising and internalising symptoms: • Externalising behaviours scores at baseline and at 12 weeks after intervention S/PCIT: Baseline 64.8; Follow up 59.0 Waiting list: Baseline 64.5; Follow up 62.9 (ES -Cohen's d=-0.38, p=0.000). • ECBI Intensity scores at baseline and at 12 weeks after intervention S/PCIT: Baseline 149.8; Follow-up 133.7 Waiting list: Baseline 149.1; Follow-up 143.1 (ES -Cohen's d=-0.27, p=0.019). • ECBI Problem scores at baseline and at 12 weeks after intervention S/PCIT: Baseline 19.1; Follow-up	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	N=61 female carers (less families than waiting list as recruitment started from 2007–9). • Sample size - 152 female carers. **Intervention category:** Parent–Child Interaction therapy. **Intervention:** Parent-Child Interaction Therapy was developed to improve parenting skills and parent–child interactions among families struggling with their children's (aged 3–7) behavior problems. It was conducted using didactic sessions to teach parents communication skills that foster positive parent–child relationships. Standard Parent–Child Interaction Therapy (S/PCIT) has two phases termed child-directed interaction (CDI) and parent-directed interaction (PDI). Progression from the first to the second phase occurred after 6–8 coaching session regardless of whether mastery criteria was achieved. Each phase begins with a didactic session designed to teach specific skills. The remainder of PCIT involves '… direct coaching sessions that provide the parent with immediate praise for appropriate responses to their child's behavior and remediation of inappropriate responses' (p257). In this study, treatment was concluded after a maximum of 12	13.5 Waiting list: Baseline 18.0; Follow-up 17.5 (ES -Cohen's d=-0.61, p=0.000). • Internalising symptoms scores at baseline and at 12 weeks after intervention S/PCIT: Baseline 54.6; Follow-up 49.8 Waiting list: Baseline 56.5; Follow-up 55.1 (ES -Cohen's d=-0.30, p=0.014). **Effect sizes - Caregiver/parent health and wellbeing outcomes:** Parental stress and depression. • Parental stress scores due to the child from baseline to 12 weeks after the intervention: S/PCIT: Baseline: 134.4; Follow-up 125.5 Waiting list: Baseline 132.5; Follow-up 130.5 (ES-Cohen's d=-0.24; p=0.041) • Parental stress scores due to the parent from baseline to 12 weeks after the intervention: S/PCIT: Baseline 147.7; Follow-up 144.7 Waiting list: Baseline 145.4; Follow-up 144.4 (ES-Cohen's d=-0.07; p=0.591) • Parental depression scores from baseline to 12 weeks after the intervention: S/PCIT: Baseline 14.0; Follow-up 12.0 Waiting list: Baseline 15.1; Follow-up 11.0 (ES-Cohen's d=0.19; p=0.153). **Narrative findings – Risk of abuse and neglect:** There were no significant differences between S/PCIT and waitlist participants in changes in total child abuse potential scores. **Narrative findings – Quality of parenting and parent-child relationships:** There was large effects observed for S/PCIT participants compared to waitlist for praise and descriptions and reflections, and medium-to-large effects in decreasing questions, commands, and negative talk. Also, a significant medium	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	sessions. The coaching intervention was delivered by six psychologists trained in PCIT who received weekly supervision. Masters and doctoral level psychologists trained in PCIT implemented the intervention. **Comparison intervention:** Waiting list 'Participants allocated to the wait-list were contacted weekly by phone by an allocated PCIT psychologist for brief conversations regarding family and other concerns. Parents in the waitlist group were asked to refrain from family therapy and therapeutic assistance with child behavior management for the duration of 12 weeks. At the end of 12 weeks, families were offered S/PCIT' (p257). **Outcomes measured:** • Risk of abuse and neglect - Parents' child abuse potential, using the Child Abuse Potential Inventory. • Quality of parenting and parent-child relationships - Parent observed behaviors using the Dyadic Parent–Child Interaction Coding System III to assess the quality of parents' verbalisations when interacting with their children. • Children and young people's health and wellbeing outcomes - Parents'	effect was observed for parental sensitivity, with greater improvement among S/PCIT participants compared to waitlist. **Narrative findings – Children and young people's health and wellbeing:** S/PCIT participants reported greater reductions in their child's externalising behaviours and internalising symptoms compared to waitlist participants with small-to-medium effects. **Narrative findings – Parental health and wellbeing:** The authors also reported that a small but significant effect for reductions in parent stress attributed to the child compared to the waitlist participants. There were no significant differences between S/PCIT and waitlist participants in changes in maternal depression and stress due to parent concerns.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	report on child externalising and internalizing symptoms using the ECBI (Child Abuse Potential Inventory). • Caregiver/parent health and wellbeing outcomes - parent stress and depression measured using the Parenting Stress Inventory, and the Beck Depression Inventory II.		

22. Thomas R and Zimmer-Gembeck MJ (2011) Accumulating evidence for parent-child interaction therapy in the prevention of child maltreatment. Child Development 82: 177–92

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The study aimed to examine the effectiveness of a standard Parent Child Interaction Therapy programme for mothers at a high risk of child maltreatment or those with a history of child maltreatment. The authors note that there is little evidence on the effectiveness of parenting programmes for families involved with the child protection system. The authors also aimed to identify which outcomes	**Participants:** • Caregivers and families. Mothers at a high risk of or with a history of child maltreatment referred by government agencies, identified as a 'suspect' by a professional, or self-referred. Mothers of children who had been sexually abused were excluded as was one mother with substance abuse issues. **Sample characteristics:** • Age - Not reported by group. The mean age of the total sample was 33.5 (SD=8.9). Mean age of participants children was 5 (SD=1.6). • Sex - All participants were female. 71% of the participants children	**Effect sizes - Incidence of abuse and neglect:** Suspected maltreatment: As the control group was offered treatment after 12 weeks data from this group was not included in the analysis. Instead, analysis (chi-square) focused on whether completion of treatment (in the intervention group) was associated with fewer notifications to child protection services. There was a significant difference between those who completed treatment and those who did not. Of those who completed treatment (n=43), 17% were the subject of notifications post treatment compared to 43% (n=53) of families who dropped-out of the programme ($x^2=7.7$, $p<0.01$). The study also found that participants who had been referred via child protection services were less likely to be the subject of further notifications post treatment. 47% of these families who completed treatment received another notification compared to 73% of families who had not completed treatment ($x^2=2.8$,	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity score:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
were associated with a reduction in child abuse and the role of maternal sensitivity specifically. **Methodology:** RCT. **Country:** Not UK - Australia. **Source of funding:** Government - Partially funded by the Future Directions Prevention and Early Intervention Trials, Queensland Department of Child Safety, Australia. No other details on funding provided.	were boys and 29% were girls. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - The study included children who had been '... physically maltreated, emotionally maltreated, and/or neglected' (p 180). Children who had been sexually abused were excluded from the study '... because sexual abuse is contraindicated for Parent Child Interaction Therapy' (p180). • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - n=51 at baseline. 71% completed a follow-up assessment. • Intervention numbers - n=99 at baseline. 42% completed all follow-up assessments. • Sample size - Comparison group: n=51 at baseline. 71% completed a	p=0.092). **Effect sizes - Risk of abuse and neglect:** Child abuse potential at 12 week assessment: Intervention -0.02 (d), comparison -0.09 (d). Between group 0.08 (d). F value: 0.22, p value: 0.644. Child abuse potential (within group) at treatment completion compared to baseline (n=41): -0.40 (d). F value 20.83, p value <0.001. **Effect sizes - Quality of parenting and parent-child relationships:** Observed parenting behaviours at 12 week assessment - Praise: Intervention 2.13 (d), control -0.05 (d). Between group 2.18 (d). F value 30.50, p value <0.001. Description and reflection: Intervention 1.17 (d), control 0.22 (d). Between group 0.95 (d). F value 14.03, p value <0.001. Questions: Intervention -1.66 (d), control -0.18 (d). Between group -1.48 (d). F value 25.49, p value <0.001. Commands: Intervention -0.54 (d), control -0.15 (d). Between group -0.39 (d). F value 3.19, p value .078. Observed maternal sensitivity at 12 week assessment - Intervention -0.06 (d), control 0.12 (d). Between group -0.18 (d). F value 0.38, p-value 0.540. Observed parenting behaviours at treatment completion compared to baseline (n=41) – Praise: 1.88 (d). F value 35.43, p value <0.001. Description and reflection: 1.26 (d). F value 31.92, p value <0.001. Questions: -1.70 (d). F value 68.34, p value <0.001. Commands: -0.63 (d). F value 10.49, p-value 0.003. Observed maternal sensitivity at treatment completion compared to baseline (n=41): 0.38 (d). F value 4.36, p value 0.044. **Effect sizes - Children and young people's health**	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	follow-up assessment. Intervention group: n=99 at baseline. 42% completed ALL follow-up assessments. **Intervention category:** Parenting programmes. **Intervention:** Standard (i.e. without a motivational component) Parent Child Interaction Therapy programme. Included 'didactic sessions' and coaching sessions (conducted with parent and child as therapist observes via a one-way mirror and communicates with the parent using a 'bug-in-the-ear'). Intervention provided by psychologists. Parents were first coached in Child Directed Interaction and when 'mastery' of this was achieved mothers were taught Parent Directed Interaction. Child Directed Interaction teaches '... specific nondirective, interactional parent skills' (p181). The aim is to foster a positive relationship in which the parents reward desirable behaviour and ignore undesirable behaviour. Parent Directed Interaction helps parents to understand their child's developmental level and manage their own expectations in relation to this and their child's behaviour. It also teaches behaviour management strategies and aims to help parents to develop reasonable and consistent limits. Parents	**and wellbeing outcomes:** Child behaviour problems (parent reported) at 12 week assessment - Externalising: Intervention -0.47 (d), control -0.07 (d). Between group -0.40 (d). F value 6.66, p value 0.12. Intensity: Intervention -0.63 (d), control 0.01 (d). Between group -0.64 (d). F-value 17.60, p value <.001. Problematic: Intervention -0.64 (d), control 0.07 (d). Between group -0.71 (d). F value 11.01, p value .001. Internalising: Intervention -0.36 (d), control -0.21 (d). Between group -0.15 (d). -value 1.41, p value .239. Child behaviour problems (teacher reported, externalising and internalising) at 12 week assessment – ds = -0.13 to 0.16. (p values and data are not reported). Child behaviour problems (parent reported) at treatment completion compared to baseline (n=41) - Externalising: -0.78 (d). F value 25.48, p-value <0.001. Intensity: -1.27 (d). F value 112.50, p value <0.001. Problematic: -1.33 (d). F value 65.34, p value <0.001. Internalising: -0.64 (d). F value 16.74, p value <0.001. **Effect sizes - Caregiver/parent health and wellbeing outcomes:** Parent stress at 12 week assessment - Stress due to child: Intervention -0.33 (d), control -0.20 (d). Between group -0.13 (d). F value 4.95, p value .029. Stress due to parent: Intervention -0.29 (d), control 0.00 (d). Between group -0.29 (d). F value 5.59, p-value .021. Parent stress at treatment completion compared to baseline (n=41) – Stress due to child: -0.83 (d). F-value 52.69, p value <0.001. Stress due to parent: -0.50 (d). F value 27.78, p value <0.001. **Narrative findings – Effectiveness:** Between group differences at 12 weeks using ANOVA: There were no statistically significant differences between groups on	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	were also expected to practice these skills at home. Treatment was completed when parents achieved 'mastery' rather than being time-limited. **Comparison intervention:** Waitlist only although parents were contacted on a weekly basis to discuss concerns related to their family or other issues. **Outcomes measured:** • Incidence of abuse and neglect - Suspected maltreatment was measured using official records from child protection services. Substantiated maltreatment after notification was not measured and no data on perpetrators was available as notifications were recorded with reference to children specifically. • Risk of abuse and neglect - Parent child abuse potential was measured using the Child Abuse Potential Inventory (Milner 1986). • Quality of parenting and parent-child relationships - Observed parenting behaviours were measured using the Dyadic Parent-Child Interaction Coding System III (Eyberg et al. 2004). Maternal sensitivity (observed) was measured using a modified version of a subscale of the Emotional Availability Scales (Biringen et al. 2000).	the child abuse potential scale; and between groups differences in notifications to child protection services were not measured. There was a significant difference between groups (in favour of the intervention group) on observed verbalisations of praise (p=<0.001), observed verbalizations of description and reflection (p=<0.001), and observed questions (p=<0.001). However, no significant differences were found between groups on observed commands or observed maternal sensitivity. The authors report that there were no significant differences between groups on teacher reported child behaviour problems (internalising and externalising), however the data on this measure are not included in the tables, which only include parentally reported child behaviour problems. Significant differences were found between groups (in favour of the intervention group) on parentally reported externalising behaviours of the child (p=0.012), however there were no significant differences between groups on parental reports of internalising problems. Significant between groups differences (in favour of the intervention) were found using the ECBI intensity scale (p=<0.001) and the ECBI problem scale (p=0.001). On measures of parental stress, statistically significant between groups differences (in favour of the intervention group) were found in parental stress 'due to the child' (p = 0.029) and parental stress 'due to the parent' (p=0.021). When between group differences were found the authors calculated clinical significance and reliable change indices.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Children and young people's health and wellbeing outcomes - The authors measured child symptoms using four scales. Child behaviour problems (intensity and whether these are deemed problematic) were measured using the Eyberg Child Behavior Inventory (reported by parents) and the Sutter-Eyberg Student Behavior Inventory-Revised (reported by teachers). (Both developed by Eyberg and Pincus 1999 and designed for children aged 2-16). Internalising and externalising symptoms were measured using the Child Behavior Checklist (reported by parents) and the CBCL/Teacher Report Form (reported by teachers). (Achenbach, 1991). The Child Behavior Checklists for children aged 4–18 and 2–3 were used. • Caregiver/parent health and wellbeing outcomes - Parent stress was measured using the Parenting Stress Inventory (Abidin 1990). **Follow-up:** Intervention and control groups were followed up at 12 weeks. After this the control group was offered treatment but data from this group was not used in further analysis. Parents randomised to the intervention group were also assessed at treatment completion and one month later.	Within group effects from baseline to treatment completion using ANOVA: There were significant declines in child abuse potential (p=< 0.001), parental reports of externalising child behaviours (p=< 0.001), parental reports of internalising child behaviours (p=<0.001), ECBI intensity scores (p=<0.001), ECBI problem scores (p=<0.001), stress due to the child (p=<0.001), stress due to the parent (p=<0.001), observed verbalisations of praise (p=<0.001), observed verbalisations of description and reflection; and observed verbalisations of questions (p=<0.001). There were no significant within group effects on measures of observed maternal sensitivity and observed commands. Differences between those who completed treatment and those who did not: The authors report significant differences in notifications to child protection services between mothers who completed treatment and those who did not (p<0.01).	

23. Zielinski DS, Eckenrode J, Olds DL (2009) Nurse home visitation and the prevention of child maltreatment: Impact on the timing of official reports. Development and Psychopathology 21: 441-453

Research aims.	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating.
Study aim: 'This study examined the effects of the Nurse Family Partnership (NFP), a program of prenatal and infancy home visiting by nurses, on the timing of verified reports of child maltreatment' (p441). This paper builds on a study reported in Olds et al. (1986) and Olds and Eckenrode (1997). **Methodology:** RCT. **Country:** Not UK – USA. **Source of funding:** • Government. • Voluntary/Charity.	**Participants:** • Children and young people - Firstborn children of predominantly unmarried, low-income mothers. • Caregivers and families - Predominantly unmarried, low-income mothers. **Sample characteristics:** • Age - 47% <19 years of age. • Sex - Apparently all female. • Ethnicity - 89% European American, 11% African American. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - n=144.	**Effect sizes - Incidence of abuse and neglect (all types of maltreatment):** Data analysed using survival functions, with estimated Cox proportional-hazard regression models. Main effect of group was not significant across the 15-year period (68% of comparison children 'survived' to age 15, compared to 76% of nurse-visited children - B=-0.31, se=0.26, p=0.23). Differences in shape of survival functions was tested by examining treatment x time interactions. Interaction between treatment group and continuous time was not significant (B=-0.07, se=0.06, p=0.26). Interaction between treatment group and time periods was significant, showing a significantly higher treatment effect between ages 4 and 15 compared to ages 0 to 4 (B=2.13, se=0.43, p=0.00). Analysis of neglect cases There was a main effect of treatment group, with nurse-visited children less likely to ever be reported for neglect (B=-0.53, se=0.29, p=0.06). There was a marginally significant interaction between group and continuous time, in favour of the treatment group (B=-0.15, se=0.08, p=0.07). There was a marginally significant interaction between group and time period, in favour of the treatment group (B=2.40, se=0.51, p=0.00). The high risk subgroup showed a significant main effect of treatment (B=-0.89, se = 0.43, p=0.04), but treatment x time interaction was not significant (B=-0.20, se=0.13, p=0.12).	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity score:** +

Research aims.	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating.
	• Intervention number - n=116. • Sample size - total sample n=237. **Intervention category:** Family Nurse Partnership. **Intervention:** Visiting by a public health nurse throughout pregnancy, until the child's second birthday. This was in addition to sensory and developmental screening for the child at 12 and 24 months, based upon which children were referred for further clinical evaluation and treatment; free transportation services for prenatal and well-child care up to child's second birthday. The goal of nurse visits is to: a) improve the outcome of pregnancy by improving prenatal health b) improve the child's health and development by improving parents' competence in early care c) improve mothers' economic self-sufficiency. The study reports that 'although nurses used detailed assessments, record-keeping forms and protocols to guide their educational activities and work with families, they also adapted the content of their home visits to the individual needs of each family. The nurses spent considerable attention to developing a close working relationship with the mother and her family' (p444).	**Narrative findings – Effectiveness:** This study looked at how nurse home visitation affected the occurrence and timing of official reports of abuse and neglect. The study found that the intervention did not have a significant impact on the overall occurrence of official reports of abuse and neglect, but it did affect the trajectory of risk, with reports of abuse and neglect for the treatment group 'levelling off' over time compared to the comparison group. However, the authors do note that this could also represent a decrease in surveillance bias as families stop receiving the home visiting intervention (p452). Looking at all forms of maltreatment, there was no difference between conditions in relation to reports of abuse and neglect over time as a whole. However, when time was divided in to two distinct period (Child age 0–4 and 4–15) there was a significant difference, in which the likelihood of reports of maltreatment was significantly less for home-visited children in the 4–15 age category than for the 0-4 age category. The authors note that 'where maltreatment did occur in the nurse-visited group [it] was confined to instances that occurred early' (p451). They note the implications for children's development - the results suggest that risk for non-home-visited children were more persistent and durable over time, therefore potentially having more negative impacts on development. Looking at neglect specifically, reports were significantly less likely in the treatment group when measuring time as a whole, or in two distinct periods. The authors note that 'when neglect was specifically exam-	

Research aims.	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating.
	Nurses scheduled to visit once every other week during pregnancy, once a week for first 6 weeks postpartum, and on a diminishing schedule until 2 years of age. Nurses completed an average of 9 visits during pregnancy (sd=3) and 23 (sd=15) from birth to child's second birthday. **Comparison intervention:** Comparison comprised three conditions comprising successive addition of the following components: i) sensory and developmental screening for the child at 12 and 24 months, based upon which children were referred for further clinical evaluation and treatment; ii) free transportation services for prenatal and well-child care up to child's second birthday iii) visiting by a public health nurse during pregnancy. **Outcomes measured:** • Incidence of abuse and neglect - Reports of maltreatment recorded in CPS reports. Study also looks at type of maltreatment. **Follow-up:** At 15 years.	ined, first-time reports ceased among children who received the intervention by approximately age 8' (p449). In the high risk subgroup (mothers who were both unmarried and had low SES at programme initiation) there was a significant effect of treatment, meaning that there were overall fewer reports of maltreatment in the high risk subgroup who received treatment, compared to those who did not. However the time by treatment interaction effect was less strong. This may be in part due to reduced statistical power due to the smaller sample size in this subgroup.	

Review question 10 – What is the impact of interventions aiming to provide early help to children and young people identified as at risk of child sexual abuse? (Prevention of occurrence)

No eligible studies found.

Review question 11 – What is the impact of interventions aiming to provide early help to children and young people identified as at risk of female genital mutilation? (Prevention of occurrence)

No eligible studies found.

Review question 12 – What is the impact of interventions aiming to provide early help to children and young people identified as at risk of forced marriage? (Prevention of occurrence)

No eligible studies found.

Review question 13 – What is the impact of interventions aiming to provide early help to children and young people identified as at risk of internal child trafficking? (Prevention of occurrence)

No eligible studies found.

Review question 14 – What aspects of professional practice and ways of working support and hinder the effective early help of children and young people at risk of child abuse and neglect?

Review question 14 – Critical appraisal tables

1. Allen SF (2007) Parents' perceptions of intervention practices in home visiting programs. Infants and Young Children 20: 266–81

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Defensible. Interviews using open-ended questions. **How well was the data collection carried out?** Appropriately. Interviews included four open-ended questions.	**Is the context clearly described?** Clear. In the context of the Help Me Grow home visiting programme. **Was the sampling carried out in an appropriate way?** Appropriate. Randomly selected from the database of all families enrolled in two Ohio based Help Me Grow programmes. **Were the methods reliable?** Reliable. A small pilot study to test the coding system and resolve	**Does the study's research question match the review question?** Yes. Parents' perceptions of assets and barriers to intervention practices. **Has the study dealt appropriately with any ethical concerns?** Yes. Passed a full review of the university institutional review board. **Were service users involved in the study?** Yes. Implicit, no consent from participants reported.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** +

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
	any differences in interpretation. Iterative process - new categories were added at each stage of the process and any disagreement discussed until consensus was reached. **Are the data 'rich'?** Rich. **Is the analysis reliable?** Reliable. 'Content analysis procedures followed a number of steps to derive valid themes from the data (Smith 2000; Weber 1990). The researcher transcribed all parents' responses grouped by respondent and by question. Two research assistants analysed the content with the researcher to provide a system of checks and balances' (pp274–5). **Are the findings convincing?** Convincing. **Are the conclusions adequate?** Adequate.	**Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Families of children who are at risk for developmental delays or maltreatment. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Home visitation. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. A home visiting programme. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** No. USA.	

2. Ayerle GM, Makowsky K, Schücking BA (2012) Key role in the prevention of child neglect and abuse in Germany: Continuous care by qualified family midwives. Midwifery 28: e529–37

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
Is a qualitative approach appro-	**Is the context clearly described?** Unclear. Few details are	**Does the study's research question match the review**	**Overall assessment of internal validity:** -

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
priate? Appropriate. The researchers aimed to gather views on the family midwife (FM) role in order to complement quantitative data measuring the effectiveness of the role. **Is the study clear in what it seeks to do?** Clear. The qualitative component has a clear research question which is to investigate the '... factors which influence support by FM for the families, such as acceptance and access from the mothers' perspective' (pe530). **How defensible/rigorous is the research design/methodology?** Somewhat defensible. The authors provide a relatively clear account of their reasons for choosing a qualitative approach and their methods of data analysis, however it is not clear why mothers in Lower Saxony were chosen for interview rather than those in Saxony-Anhalt from whom quantitative data was collected. In addition, the authors do not specify why they did not interview Family Midwives themselves. **How well was the data collec-**	provided, only that all '... but one interview were conducted in the families' homes ...' (p e532). **Was the sampling carried out in an appropriate way?** Not sure. It is not clear which families in Lower Saxony were invited to take part in interviews or why only families in this region were considered for participation in interviews rather than those in Saxony-Anhalt from whom the majority of quantitative data was collected. The characteristics of these families are not reported meaning that the extent to which the participants are representative is unclear. **Were the methods reliable?** Not sure. Although the authors discuss triangulation this relates to the decision to collect both quantitative and qualitative data. There is no indication that qualitative data was collected from more than one source. **Are the data 'rich'?** Mixed. Although diversity of perspective is discussed the level of detail provided is relatively minimal. There is no consideration of context and no real comparative element.	**question?** Yes. The qualitative component of the research was devised to '... investigate factors which influence support by FM for the families, such as acceptance and access from the mothers' perspective' (pe530). **Has the study dealt appropriately with any ethical concerns?** Yes. Informed consent was sought from mothers participating in interviews and approval from institutional ethics committees was given. **Were service users involved in the study?** No. No indication that service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. The programme aims to prevent child abuse and neglect in 'vulnerable' families. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. 'Vulnerable' families and social workers working in this field. **Is the study setting the same as**	A key limitation of the study is the failure to give any detail regarding sampling processes and the characteristics of the participants which means that it is difficult to determine the extent to which the sample is representative. **Overall assessment of external validity:** ++ **Overall validity rating:** - Key limitations of the study include a lack of detail regarding methodological concerns such as data collection and the failure to give any detail regarding sampling processes and the characteristics of the participants means that it is difficult to determine the extent to which the sample is representative.

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
tion carried out? Somewhat appropriately. Whilst the data collection process is relatively clear there is not a great deal of detail provided and it is not clear whether structured or semi structured interviews were used.	**Is the analysis reliable?** Somewhat reliable. Interviews were recorded and then transcribed 'word-to-word'. 'Deductive' categories were determined based on the objectives of the study (e.g. satisfaction with the FM and views on their collaborative role). From these, 'inductive subcategories' were generated. The authors refer to a range of research techniques which they used and describe their continuous reflection on data interpretation as 'empirical anchoring' (p e532). The authors do not report how they dealt with discrepancies. **Are the findings convincing?** Somewhat convincing. Findings are clearly presented, internally coherent and the data is generally well referenced. However, the reporting of interviews with social workers is very brief and there are no quotations from these participants. The Guideline Committee may also wish to bear in mind that the research was originally conducted in German and quotations were translated by the authors. **Are the conclusions adequate?** Somewhat adequate. The author's conclusions seem plausible and	**at least 1 of the settings covered by the guideline?** Yes. FMs visit families in their homes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Relevant to early help - the programme aims to prevent child abuse and neglect in 'vulnerable' families. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. Views on the FM role, an intervention aiming to prevent child abuse and neglect in 'vulnerable' families. **Does the study have a UK perspective?** No. German.	

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
	coherent, however the links between data and interpretation are not always clear, for example, the authors state in their conclusions that continuous care facilitated trust of the family midwife, however, this theme was not apparent in the findings section of the study. Limitations of the study are discussed but not in significant detail and the authors conclude that because the interviewers had not been involved in the care of the families this mitigated against the possibility that the mothers would give 'socially desirable' answers (pe536).		

3. Barnes J, Ball M, Meadows P et al. (2008) Nurse-Family Partnership Programme: First year pilot sites implementation in England. Pregnancy and the Post-partum Period. London: Birkbeck, University of London

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Qualitative research undertaken within the context of a mixed methods study. Qualitative approach required to elicit views about the project. **Is the study clear in what it seeks to do?** Mixed. The stated research aims do not match the sections of the report.	**Is the context clearly described?** Unclear. Insufficient information given about the characteristics of participants - does not enable contextualisation of responses. **Was the sampling carried out in an appropriate way?** Not sure. Insufficient information given regarding sampling technique.	**Does the study's research question match the review question?** Partly. The focus of the study overall is on implementation of first year pilot sites of Nurse-Family Partnership in England. Data from service users and practitioners about what makes for effective practice is one theme of the research.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity rating:** - Key methodological weaknesses in the study include: • Unclear how sampling of participants undertaken.

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
How defensible/rigorous is the research design/methodology? Somewhat defensible. There appear to be some gaps in sampling of qualitative research participants - for example there are no interviews with partners and relatives who have chosen not to be involved in the intervention. It is also not clear by what process the sample of individuals who took part in the qualitative research were selected. **How well was the data collection carried out?** Not sure/inadequately reported. The specific questions asked of qualitative research participants are not given, nor a clear account of the process by which interviews were conducted.	**Were the methods reliable?** Not sure. It is unclear what questions qualitative research participants were asked. **Are the data 'rich'?** Poor. There is often little systematic exploration of differences in opinion. **Is the analysis reliable?** Not sure/not reported. It is unclear how qualitative data have been analysed in order to derive the themes reported. **Are the findings convincing?** Somewhat convincing. **Are the conclusions adequate?** Somewhat adequate.	**Has the study dealt appropriately with any ethical concerns?** No. No discussion of ethical concerns, no evidence that ethical approval has been sought. **Were service users involved in the study?** No. Service users involved as participants only. **Is there a clear focus on the guideline topic?** Yes. Study relates to provision of early help to families at risk of abuse and neglect. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population is parents at risk of abuse and neglect (first-time low income mothers under age of 20). **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to early help. **(For views questions) Are the views and experiences reported**	• Characteristics of participants not reported, so unable to contextualise differences of experience. • Lack of clarity about research procedures, including the questions that participants were asked. • Lack of clarity about analytic procedure and how conclusions were reached.

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
		relevant to the guideline? Yes. **Does the study have a UK perspective?** Yes.	

4. Barnes J, Ball M, Meadows P et al. (2009) Nurse-Family Partnership Programme: Second year pilot sites implementation in England. The infancy period. London: Birkbeck, University of London

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is the mixed-methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and quantitative aspects of the mixed-methods question? Yes. **Is the integration of qualitative and quantitative data (or results) relevant to address the research question?** Partly. Limited integration of different data sources, although key findings synthesised in conclusions at the end of each chapter, and in conclusions section. **Is appropriate consideration given to the limitations associated with this integration, such as the divergence of qualitative and quantitative data (or results)?** No.	**Qualitative component 1:** Interviews with mothers who terminated FNP involvement (n=42). **Is the process for analysing qualitative data relevant to address the research question?** No. Study reports that interviews were conducted with 'as many as possible' of the 352 women who left the programme. Achieved sample was 42. It is unclear whether all women who left the programme were contacted, or whether only sample were contacted. **Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. No description of the characteristics of the sample, which would allow contextualisation of results.	**Does the study's research question match the review question?** Partly. Broad evaluation study with a number of research questions. 1. How can consistency of delivery and attaining fidelity to the programme model be achieved? 2. Do families receiving FNP in infancy differ in any substantial way from the population reached during pregnancy? That is, are those that drop out different from those that remain involved in the programme? 3. What factors (the family, the nurse, the site) are associated with retention/attrition of clients? How can retention be maximised? 4. How acceptable is FNP during infancy to families and to practitioners? 5. What is the extent of father involvement during infancy in FNP and how can this be maximised?	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity rating:** - Key limitations include: Lack of information regarding sampling and characteristics of achieved sample, lack of information regarding synthesis of qualitative data, lack of information regarding reliability and validity of measures used in quantitative component of study.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants? No. **Qualitative comp 2:** Structured interviews with 157 clients in receipt of FNP with purpose of assessing potential impacts for infant and family, and what clients thought of the programme during the infancy phase. **Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. **Is the process for analysing qualitative data relevant to address the research question?** Yes. **Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. Little consideration of characteristics of participants, and how this may affect responses, or of service context.	6. What are the view of children's services commissioners about FNP and what place does it have in local service plans? 7. What is the cost of delivering FNP and does this vary between sites? Questions 3, 4, and 5 were considered to be relevant to this review question (relating to aspects of professional practice and ways of working). **Has the study dealt appropriately with any ethical concerns?** No. Not reported. **Were service users involved in the study?** No. As participants only. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population is families at risk of child abuse and neglect (young, first-time mothers), and practitioners working with them. **Is the study setting the same as**	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants? No. **Quantitative component A (including incidence or prevalence study without comparison group; case series or case report):** Structured questionnaires with family nurses. **Is the sampling strategy relevant to address the quantitative research question (quantitative aspect of the mixed-methods question)?** Yes. **Is the sample representative of the population under study?** Yes. **Are measurements appropriate (clear origin, or validity known, or standard instrument)?** Partly. Bespoke questions relating to specific aspects of the programme. Reliability and validity of the measures not reported. **Is there an acceptable response rate (60% or above)?** Yes.	**at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to early help. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes.	

5. Brand T and Jungmann T (2014) Participant characteristics and process variables predict attrition from a home-based early intervention program. Early Childhood Research Quarterly 29: 155–67

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Objectives of study clearly stated? Yes. Objective was to investigate factors predicting attrition from a home visiting programme, based on the Nurse Family Partnership. **Clearly specified and appropriate research design?** Yes. Correlational design used, examining relationship between variables and drop-out rates. As study is examining drop-out from the intervention, a comparison group not receiving the intervention would not have been appropriate. **Subjects recruited in acceptable way?** Partly. Process for recruiting participants described (either self-referral or referral by social or health services, such as gynaecologists etc.) However, no information given on how referrals were screened to ensure eligibility. **Sample representative of defined population?** No. Population for this study is all those receiving the Pro Kind programme. It is unclear to what extent the sample were representative of this population.	**Measurements and outcomes clear?** Yes. Process variables measured via: - • Baseline interview to gather information about referral source, week of pregnancy and receipt of standard prenatal care. • Home visitor encounter forms to gather information about completion of the visit, and a rating of mother's engagement (scale of 1–4). • Structured telephone interviews with participants to measure satisfaction with service (scale of 1–4) and quality of helping relationship (measured via 5-item author constructed scale). **Measurements valid?** Partly. Some process variables measured using scales/measures constructed by the authors. No indication of reliability or validity of these scales. **Setting for data collection justified?** Yes. **Were all important outcomes and results considered?** Yes.	**Does the study's research question match the review question?** Partly. The study looks at both participant characteristics and process variables which predict attrition, for the purpose of this review question we are interesting in process variables only. **Has the study dealt appropriately with any ethical concerns?** No. Not reported. **Were service users involved in the study?** No. Service users involved as participants only. **Is there a clear focus on the guideline topic?** Yes. Study relates to early help for abuse and neglect. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children and families at risk of abuse and neglect - low-income, first-time mothers in receipt of a home-based early intervention programme, based on the Nurse Family Partnership model. **Is the study setting the same as**	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + Key study weakness: Lack of validated measures for process variables.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Tables/graphs adequately labelled and understandable? Yes. **Appropriate choice and use of statistical methods?** Partly. Association between independent variables and dependent variable measured using two types of analysis. One is described as 'univariate' but would more accurately be described as 'bivariate' (series of logistic regression calculations). The second form of analysis is multivariate logistic regression, including exploration of moderator effects of maternal age, SES, immigrant background and number of risk factors. **In-depth description of the analysis process?** Yes. **Are sufficient data presented to support the findings?** Yes. **Results discussed in relation to existing knowledge on the subject and study objectives?** Yes. **Results can be generalised?** Yes. Results should be generalisable to Family Nurse Partnership, as intervention is based on	**at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to early help. **Does the study have a UK perspective?** No. German study. However, background information indicates similar service context to the UK - i.e. there is universal support for mothers in pregnancy and soon afterwards.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	this model, and may well be generalisable to other early help home visiting interventions based on similarity of population and of intervention. **Do conclusions match findings?** Yes.		

6. Brandon M, Belderson P, Warren C et al. (2008) Analysing child deaths and serious injury through abuse and neglect: What can we learn? - A biennial analysis of serious case reviews 2003–2005. London: Department for Children, Schools and Families

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is the mixed-methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and quantitative aspects of the mixed-methods question? Yes. **Is the integration of qualitative and quantitative data (or results) relevant to address the research question?** Partly. There is some synthesis of qualitative and quantitative components, for example in Chapter 6. However, the process by which the data were integrated is not clear. **Is appropriate consideration given to the limitations associated with this integration, such**	**Qualitative component 1:** Thematic analysis of 47 SCR reports. **Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Partly. Reason for selecting these 47 for further analysis appears to be convenience sampling - these were the reports for which the full Overview Reports were available. **Is the process for analysing qualitative data relevant to address the research question?** Partly. Unclear how 'emerging themes' were identified and verified.	**Does the study's research question match the review question?** Partly. The study has a series of objectives, one of which is to 'identify any lessons for policy and practice, including examples of good practice' - this is considered to be relevant to question 14. **Has the study dealt appropriately with any ethical concerns?** No. No mention of ethical approval process, although potentially of lower concern as secondary analysis of documentary sources, rather than primary research with service users. **Were service users involved in the study?** No.	**Overall assessment of internal validity:** + Key limitations of the study are a lack of clarity with respect to the way in which thematic analysis of the sub-sample of 47 reviews was conducted, and how the findings from this analysis has been integrated with quantitative analysis (see Chapter 6). However, study strength is that there is a 100% sample of SCRs from the 2003-2005 time period. **Overall assessment of external validity:** ++ **Overall validity rating:** + Limitations in qualitative aspect of

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
as the divergence of qualitative and quantitative data (or results)? No.	**Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. Little consideration of how themes are linked or otherwise to other elements of the cases. **Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?** N/A. Documentary analysis. **Quantitative component A (including incidence or prevalence study without comparison group; case series or case report):** Collection and analysis of data from total of 161 case reviews. **Is the sampling strategy relevant to address the quantitative research question (quantitative aspect of the mixed-methods question)?** Yes. Sample comprises all 161 SCRs published between 2003 and 2005. **Is the sample representative of the population under study?** Yes.	**Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population is Serious Case Reviews about cases in which children have experienced abuse and neglect (leading to death or significant harm). **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study has information about multiple aspects of practice, including early help. **Does the study have a UK perspective?** Yes.	research methodology prevent awarding ++ to this study.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Are measurements appropriate (clear origin, or validity known, or standard instrument)? Yes. Key characteristics of children and families who are the subjects of the SCRs. **Is there an acceptable response rate (60% or above)?** Yes.		

7. Devaney J, Bunting L, Hayes D et al. (2013) Translating Learning into Action: An overview of learning arising from Case Management Reviews in Northern Ireland 2003-2008. Belfast: Queen's University Belfast

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Thematic analysis of Case Management Reviews. **Is the study clear in what it seeks to do?** Mixed. The study seeks to identify 'key themes' across the 24 Case Management Reviews. It is not specified what nature of issues could be considered within this category. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. Limited justification of analytic techniques provided.	**Is the context clearly described?** Clear. Contextual information about cases reviewed is provided. **Was the sampling carried out in an appropriate way?** Appropriate. 100% sample of all CMRs in a given time period. **Were the methods reliable?** Reliable. Learning from the reviews has been triangulated with relevant research evidence as appropriate. **Are the data 'rich'?** Mixed. There are relatively few examples given to illustrate the themes identified.	**Does the study's research question match the review question?** Partly. The study's research question is about identifying 'key learning' from Case Management Reviews. Part of this involves thematic analysis of 'key themes' which include issues relevant to aspects of professional practice. **Has the study dealt appropriately with any ethical concerns?** Yes. **Were service users involved in the study?** No. **Is there a clear focus on the**	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Study contains information relevant to guideline. **Overall validity rating:** + Overall, there is a lack of description of how thematic analysis was undertaken.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
How well was the data collection carried out? Appropriately. Analysis of CMR reports.	**Is the analysis reliable?** Somewhat reliable. It is unclear how the thematic analysis was undertaken, therefore difficult to judge reliability of analysis. **Are the findings convincing?** Somewhat convincing. Themes identified are often supported by other aspects of research literature. **Are the conclusions adequate?** Adequate.	**guideline topic?** Partly. Consideration of professional practice and ways of working forms part of the analysis conducted in the study. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population includes children and young people who, at one point, showed early signs of abuse and neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. **Does the study have a UK perspective?** Yes.	

8. Domian EW, Baggett KM, Carta JJ et al. (2010) Factors influencing mothers' abilities to engage in a comprehensive parenting intervention program. Public Health Nursing 27: 399–407

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear.	**Is the context clearly described?** Clear. As part of the My Baby and Me parenting programme.	**Does the study's research question match the review question?** Yes.	**Overall assessment of internal validity:** ++ Well conducted in data collection and analyses.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
How defensible/rigorous is the research design/methodology? Defensible. **How well was the data collection carried out?** Appropriately. Audio tape-recorded interviews with coaches, interviews conducted by primary researcher (a nurse and cultural anthropologist), introduced to mothers as part of the research team. The primary researcher observed an average of five coach intervention sessions (My Baby and Me) with each of the mothers. Also included: '… observed home visits, documented field notes of coach-mother interactions and mothers' levels of engagement with family coaches and program materials, and reviewed scheduled quantitative descriptive assessment data collected on the nine mothers and their children' (p401).	**Was the sampling carried out in an appropriate way?** Appropriate. Purposive sampling. **Were the methods reliable?** Reliable. See data collection. **Are the data 'rich'?** Rich. **Is the analysis reliable?** Reliable. 'Qualitative content analysis of all observational field notes and transcribed coach interview data occurred through a reflexive and iterative process. All data were coded line by line' (p402). Research also supported by '… triangulation methods consisting of 40 home visits by the primary researcher with coaches and selected mothers over a 12-month period; a review of family history and assessment data on both the mother and the child, and descriptive quantitative data from the parent study … Peer debriefing throughout the research process included bimonthly project team meetings' (p 402). **Are the findings convincing?** Somewhat convincing. No participation from mothers.	**Has the study dealt appropriately with any ethical concerns?** Yes. Written consent from mothers was obtained. Interviews with coaches were voluntary, and they were informed that these could be stopped '… at any time' (p401). **Were service users involved in the study?** Yes. Implicit. **Is there a clear focus on the guideline topic?** Yes. Professionals' perceptions of at-risk mothers' engagement in parenting programmes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Parenting programmes. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes.	**Overall assessment of external validity:** + **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Are the conclusions adequate? Adequate. Acknowledge limitations.	**Does the study have a UK perspective?** No. US.	

9. Easton C, Lamont L, Smith R et al. (2013) 'We should have been helped from day one': A unique perspective from children, families and practitioners. Slough: National Foundation for Educational Research

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Somewhat appropriate. Given that the research question aims to explore effective support for families in which neglect is an issue it is possible that a quantitative approach would have been more appropriate. **Is the study clear in what it seeks to do?** Clear. The study aims to explore effective responses to neglect and families and practitioners views on this. Although the policy context and challenges in responding to neglect are outlined there is no discussion of relevant literature or the theories which underpin early interventions in response to neglect. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. The authors do not provide a rationale for	**Is the context clearly described?** Not sure. Limited details on the contexts in which the interviews took place are provided, but the authors note that the majority took place face-to-face, although some were carried out over the telephone and some of the practitioner interviews in groups. Similarly, only very minimal detail is provided in relation to families and practitioners. **Was the sampling carried out in an appropriate way?** Somewhat appropriate. The sampling strategy is somewhat unclear. 'All local authorities identified families who met the criteria specified by LARC5. Generally LAs identified families through their multi-agency or early intervention managers and teams. Families were also selected based on the perceived likelihood that they would want to	**Does the study's research question match the review question?** Yes. The study's research question was: 'How do we effectively support families with different levels of need across the early intervention spectrum to engage with services within an overall framework of neglect?' (piv). **Has the study dealt appropriately with any ethical concerns?** Partly. The authors do not discuss ethical issues in detail but it is noted that families agreed to participate. The authors also note that they asked the nine local authorities involved if the '... research needed local ethical approval ...' and three reported that the research had been approved by a local ethics body. **Were service users involved in the study?** No. No indication that	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall validity rating:** - The failure to provide details regarding the analysis process and the characteristics of the sample are significant limitations.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
the use of a qualitative approach which seems problematic given that they aimed to understand effective responses to neglect. This approach seems to have been selected by default as the LARC (Local Authorities Research Consortium) is described as a '... qualitative research project ...'. No details are provided on why the nine local authorities were chosen or how the families and practitioners who took part were sampled. **How well was the data collection carried out?** Not sure/inadequately reported. No details on data collection are provided.	contribute to the research. One LA in particular invited families who they knew were in a stable situation at the time of the research to take part. One LA selected families who had been involved in a specific programme; another used the practitioner groups interviews to help identify possible families to invite to participate. LAs adopted similar approaches to identifying practitioner interviewees. Some shared the research with a local board or committee to raise awareness and invite participation; one contacted a local health commissioning group to invite GPs involvement' (p 47). **Were the methods reliable?** Not sure. No details on data collection methods are provided. **Are the data 'rich'?** Mixed. Whilst details on contexts of the data are sometimes provided and there is a good sense of diversity in perspective as well as some comparative element this is not consistent. **Is the analysis reliable?** Not sure/not reported. No details on analysis techniques are provided only that authorities '... sent their	service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. Focuses on early interventions to address neglect. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children experiencing neglect (at varying levels of severity), parents and families of these children, and practitioners working with them. **Is the study setting the same as at least 1 of the settings covered by the guideline?** No. Not reported. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Early intervention in response to neglect. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. Views and experiences data focuses on early intervention in response to neglect. **Does the study have a UK per-**	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	raw data to NFER for independent analysis. This report is based on NFER's systematic analysis of that data' (p3). However, it should also be noted that in appendix B, the authors also report that '... two authorities carried out secondary analysis. It was not clear from the responses what secondary analysis had been conducted' (p45). **Are the findings convincing?** Convincing. The findings are generally clear and coherent and are supported by quotes. **Are the conclusions adequate?** Adequate.	**spective?** Yes. Nine local authorities across England participated.	

10. Fernandez E (2004) Effective interventions to promote child and family wellness: A study of outcomes of intervention through Children's Family Centres. Child and Family Social Work 9: 91–104

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. The qualitative component of the study aimed to compare parents and professionals perceptions of services to complement quantitative data evaluating the effect of the intervention/services. **Is the study clear in what it seeks to do?** Clear. The study	**Is the context clearly described?** Unclear. Only limited information regarding the characteristics of participants or settings is provided and there is no in-depth consideration of context bias although the authors note that in two parent families, each parent was interviewed separately. They also note that if the family had more than one child over the age of	**Does the study's research question match the review question?** Yes. The qualitative component of the study aimed to compare parents and professionals perceptions of services delivered through Children's Family Centres to families at risk for child abuse and neglect.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall validity rating:** - The failure to include any detail regarding the analysis process and

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
has clear objectives and there is a good discussion of the literature and context although there is no discussion of the theories underpinning the research. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. The authors give a clear explanation for their decision to use a qualitative approach although they do not discuss their chosen sampling techniques or analysis processes. **How well was the data collection carried out?** Not sure/inadequately reported. Little detail regarding data collection is provided, only that semi structured interviews were used to prompt discussion of issues such as family functioning, case planning, worker-parent relationships, expectations, etc.	eight (the minimum age which the authors felt was appropriate to interview), the '... child who was the subject of the referral and of most concern to the parent was interviewed' (p93). **Was the sampling carried out in an appropriate way?** Not sure. No details on sampling technique are provided. **Were the methods reliable?** Reliable. **Are the data 'rich'?** Rich. The data are generally quite detailed and the study aims to compare the perceptions of workers and parents which the authors manage relatively well. Contextual information is also provided. **Is the analysis reliable?** Not sure/not reported. Process of analysis is not reported. **Are the findings convincing?** Convincing. The findings are clearly and coherently presented and are supported by extracts from the original data. **Are the conclusions adequate?** Adequate. The conclusions are plausible and coherent although	**Has the study dealt appropriately with any ethical concerns?** Partly. There is no discussion of ethical approval or informed consent processes although the authors note that a number of families declined to participate and that they only interviewed children over the age of 8. **Were service users involved in the study?** No. No indication that service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. The study focuses on services delivered through Children's Family Centres to families at risk for child abuse and neglect. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Families at risk of child abuse and neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Services were delivered in family homes and at Children's Family Centres.	only minimal details on the characteristics of participants and settings are a key limitation.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	there is no discussion of the limitations encountered.	**Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Prevention of child abuse and neglect in at risk families. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. The study compared parents and professionals perceptions of services delivered through Children's Family Centres to families at risk for child abuse and neglect. **Does the study have a UK perspective?** No. Australian.	

11. Girvin H, DePanfilis D, Daining C (2007) Predicting program completion among families enrolled in a child neglect preventive intervention. Research on Social Work Practice 17: 674–85

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Objectives of study clearly stated? Yes. The study had a clear objective which was to build a model to predict which families complete the Family Connections programme, a targeted preventive intervention delivered to families who meet child neglect criteria. **Clearly specified and appropriate research design?** Yes.	**Measurements and outcomes clear?** Yes. **Measurements valid?** Yes. The measures used had established reliability and validity although some relied on self-reported data. **Setting for data collection justified?** Unclear. Data collection process is described in another paper (DePanfilis D, Dubowitz H, 2005	**Does the study's research question match the review question?** Yes. The study aimed to build a model to predict which families complete the Family Connections programme, a targeted preventive intervention delivered to families who meet child neglect criteria. **Has the study dealt appropriately with any ethical concerns?**	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Subjects recruited in acceptable way? Unclear. Details on recruitment method are not provided. **Sample representative of defined population?** N/A.	Family connections: A program for preventing child neglect. Child Maltreatment, 10: 108–23). **Were all important outcomes and results considered?** Yes. **Tables/graphs adequately labelled and understandable?** Yes. **Appropriate choice and use of statistical methods?** Yes. **In-depth description of the analysis process?** Yes. **Are sufficient data presented to support the findings?** Yes. **Results discussed in relation to existing knowledge on the subject and study objectives?** Yes. **Results can be generalised?** Unclear. The extent to which the sample were representative is not demonstrated by the authors so it is not possible to determine if the results can be generalised. **Do conclusions match findings?** Yes.	Yes. Informed consent was given by participants and the research protocol was approved by an institutional review board and reviewed annually. **Were service users involved in the study?** No. No indication that service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. The study focuses on a targeted child neglect preventive intervention. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Families who meet child neglect risk criteria. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Some components of the programme are delivered in families homes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Relevant to early help - the study focuses on targeted prevention.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Does the study have a UK perspective? No. US study.	

12. Krysik J, LeCroy CW, Ashford JB (2008) Participants' perceptions of healthy families: A home visitation program to prevent child abuse and neglect. Children and Youth Services Review 30: 45–61

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate for views and perceptions. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Defensible. Participants selected randomly. **How well was the data collection carried out?** Appropriately. Semi-structured interviews were audio taped, transcribed, coded, and analysed. Using qualitative research methods the present researchers followed a categorical-content approach (Lieblich et al. 1998; Strauss and Corbin 1990). The face-to-face interviews ranged from 30-to-90 min in	**Is the context clearly described?** Clear. In the context of the Healthy Families America programme. **Was the sampling carried out in an appropriate way?** Appropriate. Randomised sample. **Were the methods reliable?** Reliable. **Are the data 'rich'?** Rich. **Is the analysis reliable?** Reliable. A cross-case analysis of transcribed interviews data was conducted to identify and label the themes. 'The information for each category for each case was reviewed by a research team of three individuals. This process	**Does the study's research question match the review question?** Yes. Views of families and carers on home visiting programmes to prevent child abuse and neglect. **Has the study dealt appropriately with any ethical concerns?** Yes. Approved by institutional review board (IRB), All participants who were contacted agreed to be interviewed. **Were service users involved in the study?** Yes. Implicit. **Is there a clear focus on the guideline topic?** Yes. Views of service users. **Is the study population the same as at least 1 of the groups**	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
length. Participants not paid for taking part.	was iterative beginning with the initial categories and then adding categories as data were reviewed. The research team worked together to guard against bias and produce decisions that were consistent across the cases. The researchers followed guidelines for conducting reliability checks and tests of internal and external validity. Three researchers coded the interviews and worked together to produce a systematic analysis process for the study. After a preliminary report was generated it was reviewed for cross-checking purposes by program staff' (p48). **Are the findings convincing?** Convincing. **Are the conclusions adequate?** Adequate.	**covered by the guideline?** Yes – families. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Homes in the home visiting programme. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. A home visiting programme. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** No. UK.	

13. LeCroy C W and Whitaker K (2005) Improving the Quality of Home Visitation: An Exploratory Study of Difficult Situations. Child Abuse and Neglect 29: 1003–13

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Objectives of the study clearly stated? Yes. The authors clearly state their objective which was to '... use an ecological assessment model to obtain a better understanding of difficult situations that	**Describes what was measured, how it was measured and the results?** Yes. The authors used focus groups to create the Difficult Situations Inventory on which the survey was based. 'Included were	**Does the study's research question match the review question?** Yes. The study aimed '... to use an ecological assessment model to obtain a better understanding of difficult situations	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
home visitors confront when implementing home visitation services' (p1003). **Measures for contacting non-responders?** The authors do not describe the use of measures to contact non-responders. **Research design clearly specified and appropriate?** Yes. **Clear description of context?** N/A. **References made to original work if existing tool used?** N/A. The authors devised a bespoke tool - the Difficult Situations Inventory. **Reliability and validity of new tool reported?** No. The authors do not report on reliability and validity of the Difficult Situations Inventory. **Survey population and sample frame clearly described?** Partly. The authors provide some level of detail in relation to survey respondents but do not provide details of the sampling frame used to identify the sample.	specific situations in which individuals must respond effectively to be considered 'competent' ... Furthermore, these situations need to be "problematical" to the degree that how to respond is not immediately apparent' (p1005). **Measurements valid?** N/A. The authors devised a bespoke tool - the Difficult Situations Inventory. **Measurements reliable?** N/A. The authors devised a bespoke tool - the Difficult Situations Inventory. **Measurements reproducible?** Yes. **Basic data adequately described?** No. It is unclear whether the questionnaire included the full inventory which was devised through focus groups - this paper only presents the top 15 difficult situations. **Results presented clearly, objectively and in enough detail for readers to make personal judgements?** Yes. **Results internally consistent?** Yes.	that home visitors confront when implementing home visitation services' (p1003). **Has the study dealt appropriately with any ethical concerns?** Yes. The research was approved by a review board from the Healthy Families programme and completion of the survey was voluntary. **Were service users involved in the study?** No. No indication that service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. The study focused on 'difficult situations' home visitors encounter when delivering the healthy Families Arizona programme, a targeted child abuse and neglect prevention programme. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Home visitors delivering the healthy Families Arizona programme, a targeted child abuse and neglect prevention programme.	**Overall validity rating:** - The failure to determine the reliability and validity of the Difficult Situations Inventory is a significant limitation of this study.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Representativeness of sample is described? Unclear. **Subject of study represents full spectrum of population of interest?** Unclear. **Study large enough to achieve its objectives, sample size estimates performed?** Unclear. The authors do not provide sample size estimates. **All subjects accounted for?** N/A. **All appropriate outcomes considered?** N/A. The study has a correlational design and does not evaluate effectiveness.	**Data suitable for analysis?** Yes. **Clear description of data collection methods and analysis?** Yes. Surveys were mailed to home visitors. **Methods appropriate for the data?** Yes. **Statistics correctly performed and interpreted?** Yes. **Response rate calculation provided?** No. **Methods for handling missing data described?** Unclear. The authors do not report any methods for dealing with missing data. **Difference between non-respondents and respondents described?** No. **Results discussed in relation to existing knowledge on subject and study objectives?** Partly. **Limitations of the study stated?** Yes. **Results can be generalised?** Unclear. The authors do not provide details on the sample so it is not possible to determine if they	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. The homes of families at risk of child abuse and neglect. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Early help - targeted prevention. **Does the study have a UK perspective?** No. American.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	were representative. **Appropriate attempts made to establish 'reliability' and 'validity' of analysis?** Partly. The authors provide reliability scores for the five factors which explained variance. **Conclusions justified?** Yes.		

14. Martin C, Marryat L, Miller M et al. (2011) The Evaluation of the Family Nurse Partnership Programme in Scotland: Phase 1 Report – Intake and Early Pregnancy. Edinburgh: Scottish Government

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Seeking to gauge the views of clients and providers of the service. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. Little information is given with regard to how the participants in the qualitative research were sampled, what their characteristics are and how representative these are of the intervention population as a whole.	**Is the context clearly described?** Unclear. There is insufficient information about the characteristics and context of the research participants to be able to a) assess to what extent their views are likely to be generalisable to the population receiving the intervention, b) understand differences in responses. **Was the sampling carried out in an appropriate way?** Not sure. Insufficient information given regarding how sampling of respondents was carried out. **Were the methods reliable?** Un-	**Does the study's research question match the review question?** Partly. The study has a number of research questions, not all of which are relevant to the review question. It has been included because the question: 'What factors support or inhibit the delivery of the programme?' may have relevance to aspects of professional practice or ways of working. **Has the study dealt appropriately with any ethical concerns?** Yes. Study approved by South-East Scotland Research Ethics Committee.	**Overall assessment of internal validity:** - Lack of information about sampling procedure and resulting participant sample is a serious flaw. **Overall assessment of external validity:** + Awarded + for external validity as only part of study is relevant to review question. **Overall validity rating:** - Little information is given with regard to how the participants in the qualitative research were sampled, what their characteristics are

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
How well was the data collection carried out? Somewhat appropriately.	reliable. Insufficient contextual information provided to ascertain reliability of findings. **Are the data 'rich'?** Mixed. Data relatively rich in terms of reporting detail and use of quotes as appropriate, however there is a lack of contextual detail. **Is the analysis reliable?** Somewhat reliable. Analytic procedure provided in appendix. However, the way analysis is reported does not make clear how this has been used. **Are the findings convincing?** Somewhat convincing. **Are the conclusions adequate?** Somewhat adequate. Conclusions generally match analysis.	**Were service users involved in the study?** No. Service users involved as research participants only - not involved in shaping or interpreting research. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study population is families whose children are at risk of abuse and neglect (due to young age and socio-economic circumstances of mothers). **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to early help. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes.	and how representative these are of the intervention population as a whole. Lack of information about participant characteristics (for both clients and family nurses) also makes it difficult to contextualise variation in research findings.

15. Paris R (2008) 'For the dream of being here, one sacrifices ...': Voices of immigrant mothers in a home visiting program. American Journal of Orthopsychiatry 78: 141–51

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Defensible. **How well was the data collection carried out?** Appropriately. Face-to-face interview lasting for 1 hr, conducted by a bilingual/bicultural research assistant. All interviews conducted in Spanish and in participants' homes. Respondent offered a $15 gift voucher from a local market for participation	**Is the context clearly described?** Clear. In the context of a home visitation programme called 'Visiting Moms'. **Was the sampling carried out in an appropriate way?** Appropriate - purposive sampling, from current programme participants. **Were the methods reliable?** Reliable. **Are the data 'rich'?** Rich. **Is the analysis reliable?** Reliable. Inductive and deductive approach to coding using grounded theory technique, line-by-line coding of studio-taped transcripts. Two reviewers were involved in coding and analyses in consultation with a principal investigation every 3 interviews. **Are the findings convincing?** Convincing.	**Does the study's research question match the review question?** Yes. To explore Latino mothers views and experiences of a home visiting programme. **Has the study dealt appropriately with any ethical concerns?** Yes. Informed consent in Spanish. Interviews by researchers not from the health centre where the women attended to ensure confidentiality. No coercion was used (p143). **Were service users involved in the study?** Yes. Implicit. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Immigrant mothers with children at risk of child maltreatment. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes –	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Are the conclusions adequate? Adequate.	participants' homes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Home visitation. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. Mothers from an ethnic population (Latino). **Does the study have a UK perspective?** No. US study.	

16. Self-Brown S, Frederick K, Binder S (2011) Examining the need for cultural adaptations to an evidence-based parent training program targeting the prevention of child maltreatment. Children and Youth Services Review 33: 1166–72

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Defensible. **How well was the data collection carried out?** Appropriately - semi-structured interviews via telephone were audio-recorded. The	**Is the context clearly described?** Clear. In the context of SafeCare, an evidence-based, behavioural parent training programme. **Was the sampling carried out in an appropriate way?** Appropriate. SafeCare trainers identified SafeCare providers to take part in study. **Were the methods reliable?**	**Does the study's research question match the review question?** Yes. Provider's views on how to engage families in parenting programmes. **Has the study dealt appropriately with any ethical concerns?** Yes. Approved by the Georgia State University Institutional Review Board. Consent agreement from participants. **Were service users involved in**	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
interviews lasted 21–77 mins in duration. Majority of participants focused on Latino families. Each participant was compensated $50 for their time.	Reliable. Interview schedule developed with input from expert consultant in cultural competency. Interview data transcribed verbatim, independently coded and any discrepancies discussed and agreed on consensus. **Are the data 'rich'?** Rich. **Is the analysis reliable?** Reliable. **Are the findings convincing?** Somewhat convincing - no participation from families. **Are the conclusions adequate?** Adequate - acknowledge limitations.	**the study?** Yes. Implicit. **Is there a clear focus on the guideline topic?** Yes. Views and experiences of providers. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes - programme providers. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes – participant's homes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Parenting programme. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** No. US study.	

17. Stevens J, Ammerman RT, Putnam FW (2005) Facilitators and barriers to engagement in home visitation: A qualitative analysis of maternal, provider, and supervisor data. Journal of Aggression, Maltreatment and Trauma 11: 75–93

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate.	**Is the context clearly described?**	**Does the study's research question match the review**	**Overall assessment of internal validity:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is the study clear in what it seeks to do? Clear. **How defensible/rigorous is the research design/methodology?** Defensible. **How well was the data collection carried out?** Appropriately. Focus groups with mothers were held in a conference room of a downtown community agency observed by the moderator and her assistant each session. Mothers were offered transportation to the community agency and given a $50 gift certificate to a local grocery store upon completion. Supervisor and the home visitor focus groups held at a professional focus group facility, observed by the first author behind a one-way mirror. Supervisors and home visitors were not financially compensated for their time but were provided with free food and beverages. Each session (total 5) lasted approximately 2 hours and was audiotaped and transcribed.	Clear. In the context of a voluntary home visitation program in the Midwest called Every Child Succeeds. **Was the sampling carried out in an appropriate way?** Appropriate. Mothers recruited from birth hospitals, social service agencies, and community clinics. Social workers, child development specialists, and paraprofessionals from 8 different community-based social service agencies using the Healthy Families America model served as home visitors. **Were the methods reliable?** Somewhat reliable - focus groups. **Are the data 'rich'?** Rich. **Is the analysis reliable?** Somewhat reliable. Data analysis focused on group transcripts to compare and contrast barriers and reasons for participation, both across levels of participants (i.e., Mother versus Home Visitor versus Supervisor) and within levels of participants (e.g., Home Visitor High Engagement Group versus Home Visitor Moderate Engagement Group).	**question?** Yes. Facilitators and barriers to engagement in home visitation. **Has the study dealt appropriately with any ethical concerns?** No - not reported. **Were service users involved in the study?** Yes. Implicit. **Is there a clear focus on the guideline topic?** Yes. Mothers' views of home visiting programmes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. A home visiting programme. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. Engagement in home visitation.	The study is quite dated as it was published in 2005. **Overall assessment of external validity:** + **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Are the findings convincing? Somewhat convincing. **Are the conclusions adequate?** Somewhat adequate.	**Does the study have a UK perspective?** No. US study.	

18. Voice Of Young People In Care (2014) Independent Inquiry on Child Sexual Exploitation in Northern Ireland: Consultation with care experienced young people. Belfast: Voice of Young People in Care

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. The study aimed to determine the views of care experienced children and young people in relation to responses to risk of child sexual exploitation. **Is the study clear in what it seeks to do?** Mixed. The study is clear in its objectives but there is no discussion of relevant literature and very little information regarding context provided. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. The design is appropriate to the research question but the authors do not provide any rationale for the methods used.	**Is the context clearly described?** Unclear. Few details regarding participant and setting characteristics are provided, only the age, gender and placement type of the children and young people. There is no consideration of context bias although it is noted that young people participated via workshops run on a Health and Social Care Trust level, advocacy sessions run at children's homes and through questionnaires. **Was the sampling carried out in an appropriate way?** Not sure. No details on sampling methods are provided. **Were the methods reliable?** Somewhat reliable. Very little detail is provided but it is noted that young people participated via workshops run on a Health and	**Does the study's research question match the review question?** Yes. The study aimed to gather the views of care experienced children and young people regarding responses to risk of child sexual exploitation. **Has the study dealt appropriately with any ethical concerns?** No. The authors do not report approval of the research by an ethics committee or an informed consent process. **Were service users involved in the study?** No. No indication that care experienced children or young people were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. The study	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall validity rating:** - The study is very unclear on a number of key methodological issues such as sampling techniques, analysis procedures and the contexts in which data were collected.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
How well was the data collection carried out? Not sure/inadequately reported. There is very little detail provided regarding the data collection process, only that children and young people '... were supported to participate in a number of ways including at Trust wide workshops, during visiting advocacy sessions in children's homes or through questionnaires' (p3).	Social Care Trust level, advocacy sessions run at children's homes and through questionnaires. **Are the data 'rich'?** Mixed. Only limited information regarding the contexts of the data are provided and there is no real comparative element. In addition, there are not a great deal of quotations to support the authors interpretations which is disappointing considering this study is described as a consultation with care experienced young people. **Is the analysis reliable?** Not sure/not reported. No details are provided regarding the analysis process. **Are the findings convincing?** Somewhat convincing. The findings are clear and coherent but there are very few extracts from the original data. **Are the conclusions adequate?** Somewhat adequate. The conclusions are generally plausible and coherent and the authors make a series of recommendations on the basis of these however there is no discussion of any limitations encountered.	focuses on care experienced children and young people's views regarding responses to risk of child sexual exploitation. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Care experienced children and young people considered to be at risk of child sexual exploitation. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. The children and young people consulted were living in a variety of settings including kinship care placements, 'at home', or in children's homes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Early help - responses to risk of child sexual exploitation. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. The study gathered the views of care experienced children and young people regarding responses to risk of child sexual exploitation.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Does the study have a UK perspective? Yes. Northern Ireland.	

19. Woodman J, Gilbert R, Allister J et al. (2013) Responses to concerns about child maltreatment: A qualitative study of GPs in England. BMJ Open 3: e003894

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. The study aimed to explore how GPs '... understood and responded to child maltreatment-related concerns in their daily practice' (p2). **Is the study clear in what it seeks to do?** Clear. The authors have a clear objective which is to '... contribute to the scant research literature on how GPs in England can respond to maltreatment related concerns ...' (p2). They note that the role of the GP is usually conceptualised narrowly; either as a participant in social care processes or at the identification and recognition stage. **How defensible/rigorous is the research design/methodology?** Defensible. Although there is no rationale given for using a qualitative approach there is a clear description of the sampling approach	**Is the context clearly described?** Clear. There is a reasonable level of detail provided with regards to participant and settings. **Was the sampling carried out in an appropriate way?** Appropriate. There is a relatively clear description of the sampling method. The four practices were chosen in order to represent a 'geographical spread' across England and on the basis of expertise in child protection (e.g. delivered training, contributed to policy, etc.) which may not directly align with the objective of understanding current responses to child maltreatment. **Were the methods reliable?** Reliable. **Are the data 'rich'?** Rich. The data are rich and detailed but	**Does the study's research question match the review question?** Yes. The study aimed to explore '... how a small sample of GPs understood and responded to child maltreatment-related concerns in their daily practice ...' (p2). **Has the study dealt appropriately with any ethical concerns?** Yes. The study was approved by the Central London 1 NHS Research Ethics Committee. **Were service users involved in the study?** No. No indication that service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. GP responses to families where there are maltreatment concerns. **Is the study population the**	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
and the authors note the limitations associated with this. **How well was the data collection carried out?** Appropriately. There is a clear description of the data collection process. Interviews were conducted in person and '... the researcher elicited narratives by asking the participants to choose two or three "children, young people or families who had prompted maltreatment-related concerns' and describe their concerns and involvement"' (p2). These were recorded (audio only) and later transcribed.	there is no real comparative element. **Is the analysis reliable?** Somewhat reliable. The authors report the use of NVivo to analyse the data thematically using '... an inductive and interpretive approach ... paying particular attention to data that did not fit and using reflections on these instances' (pp2–3). There was some double-coding and the authors note that the '... wider research team probed and questioned interpretation throughout the study ...' (p3). however the majority of coding was completed by the researcher who had interviewed participants. Practitioners were given the opportunity to comment on preliminary results and their comments were incorporated into the final output. **Are the findings convincing?** Somewhat convincing. The findings are generally clear and coherent however they are not always supported by extracts. In addition the findings in relation to health visitors are not particularly clear, whilst no data from practice nurses is included in the study. **Are the conclusions adequate?**	**same as at least 1 of the groups covered by the guideline?** Yes. GPs working with families for whom there are child maltreatment concerns. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. GP practices. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Relevant to early help. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. The study focuses on GPs views regarding their work with families for whom there are child maltreatment concerns. **Does the study have a UK perspective?** Yes. The study focused on GPs, practice nurses and health visitors from English GP practices.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Adequate.		

Review question 14 – Findings tables

1. Allen SF (2007) Parents' perceptions of intervention practices in home visiting programs. Infants and Young Children 20: 266–81

Research aims	**Study aim:** To explore '… parents' perceptions of assets and barriers to home visitation intervention practices that are effective in meeting the needs of families of children who are at risk for developing delays or maltreatment' (p266). **Methodology:** Qualitative. **Country:** USA. **Source of funding:** Other - Brody Institute for Parent-Child Studies Dissertation Award through the Mandel School of Applied Social Sciences, Case Western Reserve University.
PICO (population, intervention, comparison, outcomes)	**Participants:** • Caregivers and families - Parents who had been involved with one of two Ohio based Help Me Grow programmes for at least three months. **Sample characteristics:** • Age - Participants were between the ages of 18–32 years. Mean age was 25 years. • Sex - Mainly mothers. • Ethnicity - African American: 35.6% White: 61.1% Others: 3.3% • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Income in US dollars <6000 22.4%; 6000-11,999: 21.5%; 12,000-23,999:39.1%; 24,000-36000: 12.1%; > 36,000: 4.8%. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:**

	• Comparison numbers - Not relevant. • Intervention number - Not relevant. • Sample size n=90. **Intervention:** The Help Me Grow home visiting programme is comprised of a) service coordination, emotional support, material support, information support, instrumental support, advocacy, service integration and promotion of informal support networks; b) family supports and parental education, parenting support and instruction; c) child development assessment and monitoring (see Table 1, p268).
Findings	**Narrative findings – qualitative and views and experiences:** Study findings are arranged according to the following main themes: **A. Parent-home visitor relationship:** Parents viewed the close bonds between themselves and home visitors as an asset: 'Can talk to her about anything at anytime, like a big sister' (p276). They also reported that the interactions they had with home visitors helped them to develop their independence and confidence. However, difficulties sometimes arose when one service co-ordinator left and parents had to form relationships with a new professional. **B. Parent-home visitor communication – Assets:** Home visitor listens or they talk together; home visitor responds to questions, gives advice; home visitor shows consideration; home visitor calls or visits regularly. (Table 4) 'She interacts with us, talks, guides, leads us in the right path.' 'Always there when I have a question; I can always call her.' Barriers: Parent desires more contact. 'They need to come more—two times a month instead of one time a month – and call to check in between' (p276). **C. Community services coordinated by the home visitor (Table 4) - Assets:** Home visitor meets general family needs; home visitor provides material resources for the children and parents; home visitor provides information or support for child needs and family needs: 'I didn't have enough money to buy a crib for my baby when we moved into the apartment so they bought me a play pen so he could sleep securely. Also brought a high chair, swing to soothe him, safety tub seat. Pampers, and wipes' (p 278). 'She provides ways for me to get to the doctor... and makes sure I get to my appointments.' 'She knows all the ins and outs of community agencies so when there was a need she'd know who to talk to.' (p278). Barriers: Inadequacy in programme resources; in service coordinator knowledge; in referral services to other agencies; in community resources. 'I would like her to have more resources. They are limited' (p278). **D. Parent education Assets:** Home visitor provides child development assessment and monitoring and parenting support and instruction 'She helps me teach him stuff and brings him books. - Learning how to roll a ball and how for me to deal with his time-outs' (p278). Barriers: Parents did not identify any barriers in this domain.

	Parents appreciated close, caring, supportive relationships with home visitors and ready accessibility of them to answer their questions and offer support. Important assets perceived by parents were the personal quality of the home visitor to listen and show consideration; and the home visitor's knowledge of community services and efforts to link families with appropriate agencies or informal support for children and the families in times of crisis. Parents also valued the home visitor's help in answering their parenting questions, and teaching them techniques to help their children. Parents felt it important to minimise transfers and found it particularly difficult to adjust to transfers between service coordinators. They would appreciate more frequent contact with their home visitor and were concerned with lack of resources which contributed to difficulties in meeting their needs.
Overall validity rating.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** +

2. Ayerle G M, Makowsky K, Schücking B A (2012) Key role in the prevention of child neglect and abuse in Germany: Continuous care by qualified family midwives. Midwifery 28: e529–37

Research aims	**Study aim:** 'To investigate factors which influence support by Family Midwives for the families, such as acceptance and access from the mothers' perspective' (pe530). The study took place in the Saxony Anhalt and Lower Saxony regions of Germany. The full study included a quantitative component taking before-and-after measures of participants in the intervention. This does not meet our evidence criteria and so has not been examined. The data presented here are from qualitative interviews with 14 mothers (conducted in Lower Saxony only). **Methodology:** Qualitative study. Part of a mixed methods study but only qualitative data has been extracted for question 14. **Country:** Germany. **Source of funding:** Government - German Federal Ministry for Family Affairs, Senior Citizens, Women and Youth - 'Early Prevention and Intervention for Parents and Children and Social Warning Systems' framework.
PICO (population, intervention, comparison, outcomes)	**Participants:** • Caregivers and families - 'Vulnerable' families. • Professionals/practitioners - Social workers with district level responsibilities and those working with individual families.

	Sample characteristics: • Age - Not reported for mothers or social workers with whom interviews were conducted. • Sex - Interviews were conducted with mothers and social workers (gender not reported). • Ethnicity - Not reported for mothers or social workers with whom interviews were conducted. • Religion/belief - Not reported for mothers or social workers with whom interviews were conducted. • Disability - Not reported for mothers or social workers with whom interviews were conducted. • Long term health condition - Not reported for mothers or social workers with whom interviews were conducted. • Sexual orientation - Not reported for mothers or social workers with whom interviews were conducted. • Socioeconomic position - Not reported for mothers or social workers with whom interviews were conducted. • Type of abuse - Not reported for mothers with whom interviews were conducted. • Looked after or adopted status - Not reported for mothers with whom interviews were conducted. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported for mothers or social workers with whom interviews were conducted. **Sample size:** • Sample size - 14 mothers from 'vulnerable' families participated in interviews. Three social workers with district level responsibilities and 3 working with individual families were also interviewed. **Intervention category:** Home visiting. **Intervention:** Family midwife. A community based professional who visits 'vulnerable' families in their home to provide advice and support on subjects such as maternal and child healthcare and nutrition, and the mother-child relationship. They also provide 'psychosocial' support and counselling and help parents to fill out forms and access other services. Visits begin in the antenatal period and continue up to the child's first birthday. The frequency of visits and the type of support vary according to each family's needs but the support is '... geared towards early prevention of child abuse and neglect' (pe529).
Findings	**Narrative findings – qualitative and views and experiences:** NB Quotations were translated from German by the authors. The researchers report on three main themes which were discussed with mothers in relation to the family midwife: **1. 'Acceptance of the Family Midwife' (p e534):** Mothers were reported to prefer that the service should commence during pregnancy rather than at the post-partum stage. The authors suggest that the physical care provided to mothers and their child as well as the psychosocial support delivered by the midwife enabled the mothers to build confidence. The authors also report that a '… less complicated …' transition between the 'caseload' midwife and the family midwife led to higher levels of trust in the family midwife, which in turn meant that mothers were able to rely on the family mid-wife as a source of support.

	Participants were also reported to value the family midwife as someone they could confide in and talk through their problems with, for example regarding the father of their child and some mothers viewed the family midwife as a friend (pe534). 'Yes, she was very much of a frie [sic] … well, so to speak, her relationship to families was very friendly, especially to me, to us (…) you could feel that, because she was happy when something worked out well- that was what I liked best with her' (pe534–5). The authors conclude that these positive views of the family midwife encouraged the mothers to ask their family midwife for specific advice. 'Really, the FM visits with you so that you can ask her any questions, or in any situation you need help, or generally questions about the baby. You can really pump her for information.' However, the authors also note that some mothers did not view their family midwife as positively, noting that families who had had a '… bad experience …' with Youth Welfare Services were sceptical towards the programme in the beginning. One participant reported that she had been concerned regarding the presence of the family midwife in her home: 'Authorities, they always think they have helped you, but I feel they have strained me more than they have been helpful. (…) I felt comfortable with the family midwife (…) in the beginning I was afraid she would walk through my apartment like the other one (…) controlling my refrigerator (…) [who] had said to me: where is your flour (…)? your rice (…)? Anyway, the baby ate nothing at the time but breast milk. That was absolutely annoying to have her walk through the apartment. And she (name of FM) was never in the other room, really, she was only here (…)' (pe535). **2. 'Access to the Family Midwife' (p e535):** The authors highlight the importance of access to the family midwife, even to families for whom a goal was to develop social networks and enhance maternal self-sufficiency, and note that the ability to contact the midwife via phone was particularly welcomed by mothers. 'We had a date (…) spontaneously (laughing), we met spontaneously, whenever she was free. I was hardly able to work, at that time. And eventually, she was really present whenever I needed her, via phone, and via text messages' (pe535) **3. 'Collaboration among providers' (p e535):** The authors note that the issue of inter-professional working, and collaboration with Youth Welfare Services in particular, was discussed in interviews with social workers (3 with responsibility for whole districts and three working directly with families). The author's report that this collaboration '… seemed

	to be efficient in terms of support for the family…', e.g. when the mother was herself a minor; and that participants viewed this way of working positively (pe535). 'I was involved in a programme for jobless minors (…) I had to do whatever I was able to. They really helped me a lot. They exchanged information and told each other what I needed (…) and that was rather good' (pe535). In their findings section, the authors also state that continuous care enabled the mothers to trust the family midwife.
Overall validity rating	**Overall assessment of internal validity:** - A key limitation of the study is the failure to give any detail regarding sampling processes and the characteristics of the participants which means that it is difficult to determine the extent to which the sample is representative. **Overall assessment of external validity:** ++ **Overall validity rating:** - Key limitations of the study include a lack of detail regarding methodological concerns such as data collection and the failure to give any detail regarding sampling processes and the characteristics of the participants means that it is difficult to determine the extent to which the sample is representative.

3. Barnes J, Ball M, Meadows P et al. (2008) Nurse-Family Partnership Programme: First year pilot sites implementation in England. Pregnancy and the Post-partum Period. London: Birkbeck, University of London

Research aims	**Study aim:** The stated aims of the evaluation are to: 'Document, analyse and interpret the feasibility of implementing the Nurse-Family Partnership (NFP) model of home visiting in ten demonstration sites in England, to determine the most effective method of presenting the model, to estimate the cost of presenting the NFP model, to determine the short-term impact on practitioners, the wider service community and the children and families, and to set the groundwork for a possible long term experimental assessment of the programme and its impacts' (p21). For the purpose of this review, we were particularly interested in the sections of the evaluation looking at 'Is the FNP acceptable in England?' (Section 5) and 'Nature of the work and best practice' (Section 8), as potentially giving useful information about aspects of professional practice and ways of working that help and hinder effective early help. These sections are both based on qualitative research with service users, families and practitioners. Therefore, only the qualitative elements of this study have been critically appraised. **Methodology:** Qualitative study. Overall, the study used a mixed methods design. However, the sections of interest to this review question are based on: semi-structured interviews with family nurses (n=47) and their supervisors (n=10); a

	sample of enrolled clients (n=106) - interviews with relatives of clients (n=44), of whom most were partners (n=30), the remainder were mothers of clients; interviews with a sample of clients who left the programme (n=20). **Country:** UK.
PICO (population, intervention, comparison, outcomes)	**Participants:** • Caregivers and families - The information reviewed here is taken from: - interviews with a sample of enrolled clients (n=106) - interviews with relatives of clients (n=44), of whom most were partners (n=30), the remainder were mothers of clients - interviews with a sample of clients who left the programme (n=20). Families involved in the programme are those in which the mother is under the age of 20 and having her first child. • Professionals/practitioners - The information reviewed here is taken from: - interviews with family nurses (n=47) and their supervisors (n=10). **Sample characteristics:** • Age - All mothers under age of 20 at time of enrolment. Age characteristics of family nurses and supervisors not given. • Sex - Service users: All primary clients are women. A sample of partners was interviewed (n=30) - it is unclear if all were male. A sample of mothers of the client (n=14) were also interviewed. Gender of family nurses and supervisors not given. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** The information reviewed here is taken from: - interviews with clients and their families (n=170) - interviews with family nurses and their supervisors (n=57). Total sample on which this review is based = 227. **Intervention:** Programme is designed for low-income first-time mothers, starting during second trimester of pregnancy. The programme has three goals: 'To improve the outcomes of pregnancy by helping women improve their prenatal health To improve the child's health and development by helping parents to provide more sensitive and competent care of the child - To improve the parental life course by helping parents plan future pregnancies, complete their education and find work' (p19).

	Visits begin at 14–16 weeks gestation - nurses visit weekly for the first month and then every other work until birth. Visits are weekly for first 6 weeks after birth, then every other week until child is 1. Visits are every other week until the child is 20 months, and then monthly until 2 years old. The programme: - Addresses modifiable risk for poor birth outcomes and child neurodevelopment impairment - After birth, focuses on developing sensitive competent care of the child, avoiding abuse, neglect and injuries - Supports mothers to gain educational qualifications, plan subsequent pregnancies and plan for employment.
Findings.	**Narrative findings – qualitative and views and experiences:** Relevant report sections: 1. Section 5 - Is the NFP acceptable in England? **A - Acceptability of the service to young pregnant women:** Reasons for accepting the service. Study reports that the majority of clients interviewed reported that 'they had been offered the support because of their age and because it was their first baby' (p50). Acceptance was based on: gaining extra information, to be part of a research project, based on a need for support e.g. 'I was pleased, I needed someone to be there for me, to talk to' (p50). This was not precluded by existing support from family members. The study reports that some young women were 'circumspect' about the offer of FNP, but willing to try. Why continue with the service? The study reports that continuation with the service was 'influenced to a great extent by clients' perception of the Family nurses, which were overwhelmingly positive' (p51). The study also reports that 'Many comments were made about the fact that the Family nurses spent a good amount of time with [families], sufficient for them to ask questions and go over information, much of which they had received via midwifery visits, but in a way they could fully understand' (p51). The study reports that clients thought that family nurses interacted with them in a different way, and were not judgemental about their being pregnant. 'I've just changed doctors and I think they're really rude there. Because [sic] where I'm young and pregnant they pick at that all the time - they said to me last time it's like kids having kids, it made me low and made me feel upset. But (FN) she is a nice person and don't treat me like a kid - she treats me like everyone else, she don't treat me like I'm different' (p52). Family nurses were described as 'more like a friend' (p52). Their capacity to hold back and not force their point of view on mothers was also valued. The authors also report that the 'strength-based focus of the programme led to mothers feeling that they were more able to admit problems' (p53). Understanding the extent of FNP: The study found that the majority of clients understood the length and nature of the commitment involved.

	Are the materials acceptable? FNP includes numerous activities and materials including smoking and diet diaries, sheets about exercise, dental care, safe sex, contraception, labour and danger signs. The study found that nearly all the materials were recalled by at least some of the clients. The study also reports that '… a number of positive comments were made about the fact that information was discussed, not just handed out' (p54). The study reports that some of the clients reported that they had not learned anything new through the programme. One participant said 'I have not learned anything, I have lots of children in my family'. Why stop FNP? As part of the evaluation, researchers spoke to a small number of clients (n=20) who had dropped out. The study reports that the main reason given was having sufficient support and/or knowledge. Other reasons included difficulties with the paperwork, difficulty fitting in visits, the long duration of involvement, and moving out of area. The study states that 'Most mentioned that they like the Family Nurse and that it had nothing to do with her' (p57). **B - Acceptability to fathers/partners:** The study spoke to a number of partners (n=30) and reports that they were generally 'pleased to be involved, although many did not expect that they would be part of the programme' (p57). The study reports that it often took a few sessions before fathers became engaged with the activities. Some fathers were initially concerned that the family nurse's presence would be 'intrusive and possibly judgemental' (p57). The study reports that some fathers thought that the appropriate strategy during the family nurse's visits was to be at home, but not necessarily present for the whole visit. **C - Acceptability to extended family:** In some cases, mothers were interviewed rather than the client's partner. The study found that most mothers were 'not taking a very active role in the FNP visits, but were aware of the topics being covered and made themselves available to discuss the materials with their daughters' (p59). **D - Acceptability to Family nurses and supervisors:** Early experiences: No data relevant to aspects of professional practice. Later experiences: No data relevant to aspects of professional practice. Coping with attrition: No data relevant to aspects of professional practice. How do family nurses retain clients? The study asked family nurses what they did to keep clients on the programme. Family nurses reported 5 aspects of practice which helped: - Meeting emotional needs - Flexibility, in terms of changing appointments and meeting places - Information, in terms of being able to offer information in more detail - Being family nurses which meant that 'they were health professionals, but with a different name and a caring approach' (p64) - Clients wanting the best for their baby and using this as the foundation for their work. **Conclusions:** The authors conclude that: - Clients and their families were positive about FNP - They liked FNP in comparison to other services, in particular that they were supported by the service, rather than 'judged' - Men who were interviewed had not expected to be involved by health professionals, and were pleased by this - Grandmothers appreciated the source of support - Most FNs reported enjoying the role and 'the challenges it offered' (p70) 2.

	Section 8 - Nature of the work and best practice **A - Benefits of FNP for practitioners:** The family nurses involved in the study reported the following benefits of the programme for them: 'Reaching real need - Using skills - Working with a structure programme - Standing shoulder to shoulder with the client - Small signs of progress - Close relationship within the FNP team - High quality training - The scope of the work - particularly working with fathers and extended family' (p91). **B - Barriers to effective working:** The family nurses interviewed identified the following barriers to being able to do their work: - Size of caseload - Last minute cancellation of visits - Insufficient planning time clients' loss of interest after the birth - Fatigue - Presence of numbers of people during a visit - Clients who cannot read or write - Problems with supervision - Having to keep separate data - Slipping back in to the health visitor role after pregnancy - Insufficient knowledge about some matters on their caseloads - Travelling long distances - Getting expenses from the PCT - Insufficient quantities of equipment - Not being informed when client has been discharged from maternity unit (pp91–2). **C - Best practice in the FNP as identified by clients:** The study reports the following features of FNP as being identified as best practice by clients: - A preference for practical help - Appreciating the health background of the FN - Helping with housing issues - Engaging with clients with whom other professionals had not managed to engage - Maintaining relationships even during safeguarding procedures - Taking a strengths-based approach - Encouraging clients to re-engage with other agencies - Findings ways to communicate with people for whom English is not their first language - Engaging directly with fathers - Working flexibly with families to ensure that fathers are involved (p97). **Conclusions:** The authors conclude the following in relation to good practice: - Part of best practice involves making the service accessible to the vulnerable families that it is aiming to reach, for example 'spending time exploring clients' lives with them' - family nurses 'use themselves' as an element in the programme.
Overall validity rating.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity rating:** - Key methodological weaknesses in the study include: • Unclear how sampling of participants undertaken. • Characteristics of participants not reported, so unable to contextualise differences of experience. • Lack of clarity about research procedures, including the questions that participants were asked. • Lack of clarity about analytic procedure and how conclusions were reached.

4. Barnes J, Ball M, Meadows P et al. (2009) Nurse-Family Partnership Programme: Second year pilot sites implementation in England. The infancy period. London: Birkbeck, University of London

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: This evaluation has a number of questions – 1. How can consistency of delivery and attaining fidelity to the programme model be achieved? 2. Do families receiving FNP in infancy differ in any substantial way from the population reached during pregnancy? That is, are those that drop out different from those that remain involved in the programme? 3. What factors (the family, the nurse, the site) are associated with retention/attrition of clients? How can retention be maximised? 4. How acceptable is FNP during infancy to families and to practitioners? 5. What is the extent of	**Participants:** • Caregivers and families - Information relevant to our review question was provided by: Data collection with four samples of service users: -Interviews with 157 clients in receipt of FNP with purpose of assessing potential impacts for infant and family, and what clients thought of the programme during the infancy phase - Surveys completed by 98 clients in receipt of FNP to determine satisfaction with the service, service use beyond FNP and involvement of partners - Interviews with 42 former clients who had terminated FNP involvement -Detailed case studies of 9 clients. • Professionals/practitioners - Information relevant to our review question was provided by structured questionnaires with nursing staff involved in offering the service (n=40). **Sample characteristics:** • Age - All clients aged under 20 at time of enrolment. Age of practitioners not reported.	**Quantitative data - Satisfaction with services:** Data relevant to aspects of professional practice and ways of working that help and hinder early help. **Chapter 3 - Retention of clients** **A - Rates of attrition:** No relevant data. **B - Who leaves, who stays?** No relevant data - analysed according to client characteristics only. **C - Reasons for leaving (p. 32):** Where participants declined further participation, reasons given included: - Needs had been satisfied (13.4% of those who left in pregnancy, 21.2% who left in infancy) • Had sufficient knowledge and support (7.0% of those who left in pregnancy, 1.8% who left in infancy). • Changed their mind and no longer wanted FNP (5.2% of those who left in pregnancy, 2.7% who left in infancy). • Pressure from family members (7.0% of those who left in pregnancy, 1.8% who left in infancy). • Dissatisfied with the programme (5.2% of those who left in pregnancy, 2.7% who left in infancy). • Returned to work (1.7% of those who left in pregnancy, 3.1% who left in infancy). • Returned to school (0.6% of those who left in pregnancy, 3.1% who left in infancy). • Refused new Family Nurse (1.7% of those who left in pregnancy, 1.3% who left in infancy). • Receiving services from another programme (1.7%	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity rating:** - Key limitations include: Lack of information regarding sampling and characteristics of achieved sample, lack of information regarding synthesis of qualitative data, lack of information regarding reliability and validity of measures used in quantitative component of study.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
father involvement during infancy in FNP and how can this be maximised? 6. What are the view of children's services commissioners about FNP and what place does it have in local service plans? 7. What is the cost of delivering FNP and does this vary between sites? Questions 3, 4, and 5 were considered to be relevant to this review question (relating to aspects of professional practice and ways of working). **Methodology:** Mixed methods - Methods providing information relevant to our review question was provided by: - • Interviews with 42 former clients who had terminated FNP to identify reasons for termination.	• Sex - All clients were female. All partners interviewed were male. Sex of staff not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • n=157 clients in receipt of FNP (interviews). • n=98 clients in receipt of FNP (surveys). • n=42 former clients who had terminated FNP involvement (interviews). • n=9 clients (detailed case studies). • n=40 nursing staff involved in offering the service (questionnaires). **Intervention category:** Home visiting. **Intervention:** As described in Barnes	of those who left in pregnancy, 1.3% who left in infancy). • No time (2.9% of those who left in pregnancy, 0% who left in infancy). • Other reason (1.7% of those who left in pregnancy, 1.3% who left in infancy). • No reason specified (0% of those who left in pregnancy, 0.4% who left in infancy) (n = 83 for those who left in pregnancy, n=93 for those who left in infancy). The study reports that family nurses assigned the following ratings (out of 10) to factors perceived to help clients stay with FNP: • A good relationship with the family nurse (9.8). • Enjoyment of the visits (8.9). • Flexibility in timing of visits (8.8). • Sensitive use of FNP materials to meet specific client needs (8.4). • Achieving some change (7.7). • Support from family members to stay with FNP (7.6). • Recognition that FNP is needed for many challenges in their life (6.7). • Support to stay with FNP from other involved professionals (e.g. social worker). • Referrals to other professionals for specific needs (6.3). • Presence of partner at the visits (5.9). **Narrative findings:** Data relevant to aspects of professional practice and ways of working that help and hinder early help.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
• Structured interviews with 157 clients in receipt of FNP with purpose of assessing potential impacts for infant and family, and what clients thought of the programme during the infancy phase. • Brief surveys – Surveys completed by 98 clients in receipt of FNP to determine satisfaction with the service, service use beyond FNP and involvement of partners; structured questionnaires with nursing staff involved in offering the service (n=40). **Country:** UK. **Source of funding:** Government.	(2008): Programme is designed for low-income first-time mothers, starting during second trimester of pregnancy. The programme has three goals: 'To improve the outcomes of pregnancy by helping women improve their prenatal health To improve the child's health and development by helping parents to provide more sensitive and competent care of the child - To improve the parental life course by helping parents plan future pregnancies, complete their education and find work' (p19). Visits begin at 14–16 weeks gestation - nurses visit weekly for the first month and then every other work until birth. Visits are weekly for first 6 weeks after birth, then every other week until child is 1. Visits are every other week until the child is 20 months, and then monthly until 2 years old. The programme: - Addresses modifiable risk for poor birth outcomes and child neurodevelopment impairment - After birth, focuses on developing sensitive competent care of the child, avoiding abuse, neglect and injuries - Supports mothers to gain educational	• **Chapter 1 - Introduction:** No data relevant to Q14. • **Chapter 2 - Delivering FNP with fidelity:** No data relevant to Q14. • **Chapter 3 - Retention of clients** • **A - Rates of attrition:** No data relevant to Q14. • **B - Who leaves, who stays?** No data relevant to Q14 - analysed according to client characteristics only. • **C - Reasons for leaving (p32):** Study reports that common reasons for attrition were: moving out of the FNP area, many missed appointments, or the FN being unable to locate the client. Family nurses were asked about strategies that they would use if a client was intending to leave the programme. The study reports that the most commonly used strategies were: - to go to their team - to find out in more detail what particular issues of concern were for a client, often using motivational interviewing (p.34). The study reports that family nurses rated the following factors most highly in terms of help clients stay with FNP: having a good relationship with the family nurse, enjoyment of the visits, flexibility in timing of visits, sensitive use of FNP materials to meet specific client needs. • **E - Clients' thoughts on attrition:** In semi-structured interviews with clients who left the programme, they were asked 'if anything about the nurse's behaviour, the FNP materials or the frequency of the visits led them to decide to leave' (p36). The study reports that only five interviewees mention the family nurse, 4 the frequency of the visits and 3 the actual materials. The study notes that almost every client leaving the programme de-	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	qualifications, plan subsequent pregnancies and plan for employment.	scribes their family nurse positively. The study reports that '… friendly, easy to talk to and providing really helpful information were the most common themes' (p37). A small number of clients commented negatively on being asked to go to activities. 'She was OK but she was bugging me. Kept telling me to go places about my reading and writing and I did not want to. Told me to go to [mother and baby group] and I didn't want to. I felt ashamed to say No I didn't want to go. She kept texting me and bugging me' (p37). The study reports that a small number of clients commented negatively on the family nurse speaking to other professionals, which was perceived as 'breaching confidentiality'. One client also reported that a 'less intrusive' approach might have persuaded her to stay involved with the programme. The study reports that some leavers commented negatively on the family nurse's level of knowledge. The study reports that most clients who left the programme enjoyed the programme materials to some extent, though 'many described how they found them too much to take in and only used them selectively' (p. 38). 'Sometimes liked the materials, I didn't like all the paperwork' (p39).	

5. Brand T and Jungmann T (2014) Participant characteristics and process variables predict attrition from a home-based early intervention program. Early Childhood Research Quarterly 29: 155–67

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: 'This study investigated factors predicting attrition in a sample of 434 low-income, first-time mothers in a German program modeled on the Nurse-Family Partnership Program' (p155). The study investigates the impact of both participant characteristics and 'process variables' associated with delivery of the intervention with drop-out from the intervention. For the purposes of this review, data on process variables only have been extracted. **Methodology:** Cross-sectional study - Correlational design. **Country:** Germany. **Source of funding:** Government - voluntary/charity.	**Participants:** Caregivers and families - 434 low-income, first-time mothers in receipt of a home-based early intervention programme, based on the Nurse Family Partnership model. **Sample characteristics:** • Age - Mean age 21.3, standard deviation 4.3 • Sex - All participants were female • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Mean receipt of social support reported as '51.1' - unclear what this refers to. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** 434 mothers. **Intervention category:** Home visiting.	**Quantitative data - Service outcomes:** Measures of the relationship between a number of variables and likelihood of drop out. **Multivariate analysis of relationship between process variables and addressable attrition (n=362):** • Self-referral - Odds ratio 0.22; 95% Confidence interval [0.06;0.82]; p=0.025 • Week of pregnancy - OR 0.95; 95% CI [0.90; 1.01]; p=0.080 • Change in home visitor - OR 0.69; 95% CI [0.21; 2.24]; p=0.539 • External midwife - OR 0.82; 95% CI [0.40; 1.67]; p=0.579 Unsuccessful visit attempts - OR 1.06; 95% CI [1.04; 1.09]; p=0.000 • Visits with partner - OR 1.01; 95% CI [1.00; 1.02]; p=0.075 Visits with grandmother - OR 0.97; 95% CI [0.95; 1.00]; p=0.051; • Impact on early attrition OR 0.95 95% CI [0.92; 1,00]; p=0.046 • Engagement in home visits - OR 0.32; 95% [0.18; 0.57]; p=0.000 • Helping relationship - OR 0.29; 95% CI [0.08; 1.12]; p=0.070 • Satisfaction with service - OR 0.79; 95% CI [0.57; 1.10]; p=0.162 • Time spent on parenting - OR 0.95; 95% CI [0.92; 1.00]; p=0.039. • Amount of variance explained by process variables as a whole = 0.43 (Nagelkerke's R2).	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + Key study weakness: Lack of validated measures for process variables.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Intervention: Home visiting programme based on Nurse Family Partnership (Olds 2006), seeking to 'improve maternal and child health, enhance maternal lifecourse development, strengthen parenting skills and improve the informal and formal systems of social support' (p158). The programme in Germany is delivered by midwives, or pairings of midwives and social workers. All home visitors therefore hold either a university or college degree in social work, or were state-certified midwives. Home visitors also received approximately 16 days of in-service training and 1h of clinical supervision per week from a supervisor with a university degree in social work or psychology, and with additional qualifications in coaching techniques. **Outcomes measured:** • Service outcomes - Dependent variable was attrition (drop-out) from the service. The authors distinguish between 'natural' attrition, occurring for reasons not to do with the programme, and 'addressable' attrition which 'might be changed by modifications of the program model' (p156). The authors have also examined early attrition (dropping out	**Narrative findings:** The analysis showed that the following process variables were significantly associated with drop-out from the programme: - Self-referral - Unsuccessful visit attempts - Maternal engagement in the home visits - Time spent on parenting in the home visits. There was no significant association between involvement of grandmothers and overall addressable attrition, but there was an impact on 'early attrition' (before 25% of enrolment time completed). Variables which were not shown to be significantly associated with drop-out at any time were: - Week of pregnancy (no further information given about what this refers to) - Changes in home visitor - External midwife (no further information given about what this refers to) - Visits with partner - Quality of helping relationship (although this was approaching significance at p=0.070) - Satisfaction with service. The authors explain their findings as follows. For the variables which were shown to have a significant association with drop out: - Self-referral - this may reflect a higher level of motivation, following a conscious decision to engage with the programme - Unsuccessful visit attempts and low maternal engagement - both suggest a low level of interest in the programme - Time spent on parenting issues - The authors note that this suggests that focusing on the unborn child was more engaging than discussing other topics such as maternal health behaviour or relationships issues. They also note that similar findings have been reported in other studies (Roggman et al. 2008;	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	before 25 % of enrolment time completed) and late attrition (dropping out between 25% and 75% of enrolment time completed). • Independent process variables were: - Referral source, week of pregnancy, receipt of standard prenatal care (gathered via baseline interview) - Visit completion, mother engagement with visits (scale of 1-4), family involvement in visits, unsuccessful visits, programme content (gathered via home visit encounter form). • Satisfaction with service (scale of 1-4) and quality of helping relationship (measured via 5-item author constructed scale) gathered via telephone interview.	O'Brien et al. 2012). The authors also comment that the study did not specifically examine methods of delivering the intervention, so spending more time on parenting 'may not have a beneficial effect on retention unless the topic is presented in an engaging way' (p165). Participation of grandmothers was shown to be related to early attrition only - if grandmothers did not participate, mothers were more likely to drop out before completing 25% of their enrolment time. The authors interpret this as showing that the participation of grandmothers can help to strengthen mothers' commitment to the programme. They also suggest that the presence of a grandmother may be indicative of greater family cohesion. In terms of variables which were not shown to be significantly associated with drop-out: - Week of pregnancy - no explanation suggested - Changes in home visitor - no explanation suggested - External midwife - no explanation suggested - Visits with partner - The authors note that the moderator analysis suggests that 'boyfriends and husbands were only a stabilizing factor when the overall burden was high' (p164). - Quality of helping relationship - the study notes that this was 'associated with attrition in the univariate analysis but not in the multivariate analysis. This was possibly due to their correlation with maternal engagement during the visits' (p165). Limitations: The authors note the limitation imposed by the correlational design of the study, which restricts the ability to draw causal inferences from the findings.	

6. Brandon M, Belderson P, Warren C et al. (2008) Analysing child deaths and serious injury through abuse and neglect: What can we learn? - A biennial analysis of serious case reviews 2003–2005. London: Department for Children, Schools and Families

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The aims of the study are: 'i. To provide descriptive statistics from the agreed full sample (i.e. 161 cases), illustrated by some examples from the reviews ... ii. To scrutinise a sub sample if cases (i.e. 47) to chart thresholds of multi-agency intervention at the levels specific in Every Child Matters (Cm 5860 2003) ... iii. Building on the learning from the first two objectives, to seek a meaningful analysis by identifying some ecological-transactional factors within the sub-sample of reviews ... iv. To provide practice tools for use by Local Safeguarding Children Boards and practitioners and to identify any lessons for	**Participants:** Children and young people - Sample comprises: 161 Serious Case Review reports, conducted 'when abuse and neglect are known or suspected factors when a child dies (or is seriously injured or harmed), and there are lessons to be learnt about inter-agency working to protect children' (p7). The 161 SCRs studied were notified during the period April 2003 to March 2005. **Sample characteristics:** • Age - The ages of the children who were subject to SCRs considered in the study were as follows: 0-1 month - 13% 2–3 months - 19% 4-6 months - 11% 7–12 months - 4% 1-3 years - 18% 4–5 years - 2% 6-10 years - 7% 11–15 years - 16% 16 years + - 9% (n not given - assume 161). • Sex - Female - 55% Male - 45% (n not given - assume 161). • Ethnicity - White/White British (74%) Mixed (6%) Black/Black British (13%) Asian/Asian British (6%) Other ethnic group (1%) (n=136). • Religion/belief - Not reported. • Disability - Disability recorded in	**Narrative findings – qualitative and views and experiences:** **Implications for services Universal services and early needs Levels 1 and 2 (p102).** The report identifies the following implications for universal services and early needs: - • Practitioners needs to have a holistic understanding of children and families and awareness of how factors may interact to increase risk in the family. • Staff working with children with additional needs should understand that they are working within the safeguarding continuum, and not in a separate sphere of activity. • Practitioners should be aware of common causes of child injury and death in their work with parents, including 'loss of control and volatility', overlying and water scald (suggest that the detail of these is more relevant to NCCSC questions relating to recognition).	**Overall assessment of internal validity:** + Key limitations of the study are a lack of clarity with respect to the way in which thematic analysis of the sub-sample of 47 reviews was conducted, and how the findings from this analysis has been integrated with quantitative analysis (see Chapter 6). However, study strength is that there is a 100% sample of SCRs from the 2003–5 time period. **Overall assessment of external validity:** ++ **Overall validity rating:** + Limitations in qualitative aspect of research methodology prevent

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
policy and practice, including examples of good practice' (p15). **Methodology:** Other - Analysis of Serious Case Reviews - analogous to thematic analysis of multiple case studies - therefore it is most appropriate to appraise this study using a qualitative study critical appraisal tool. **Country:** UK. **Source of funding:** Government.	5% of cases (n=161). • Long term health condition - Information on long term health conditions available for 'intensive' sample only (n=47). Complex health needs - 9%. Chronic illness - 11%. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Head injury - 16% Sudden Infant Death < 4% Overlying - 4% Physical assault - 35% Neglect - 21% Poisoning/overdose - 4% Suicide - 9% Sexual abuse - 4% Gone missing - 4%, other <4% (n=161). • Looked after or adopted status - In care at time of incident - 10% (n=159). • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Main sample n=161 SCR reports. • Intensive sample n=47 SCR reports.		awarding ++ to this study.

7. Devaney J, Bunting L, Hayes D et al. (2013) Translating Learning into Action: An overview of learning arising from Case Management Reviews in Northern Ireland 2003-2008. Belfast: Queen's University Belfast

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The aim of	**Participants:**	**Narrative findings:** Key themes identified through	**Overall assessment**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
the report is to '… present key learning from the first 24 case management reviews commissioned and completed [in Northern Ireland] between the commencement of the current process for case management reviews in 2003, up until the end of 2008' (p17). **Methodology:** Other - Analysis of Serious Case Reviews - analogous to thematic analysis of multiple case studies - therefore use qualitative study critical appraisal tool. **Country:** UK. **Source of funding:** Government.	• Children and young people - Case Management Reviews concerning children and young people who have died or been seriously injured, and abuse or neglect is known or suspected to have been a contributing factor. • Caregivers and families - Case Management Reviews concerning children and young people who have died or been seriously injured, and abuse or neglect is known or suspected to have been a contributing factor. • Professionals/practitioners - Case Management Reviews concerning children and young people who have died or been seriously injured, and abuse or neglect is known or suspected to have been a contributing factor. **Sample characteristics:** • Age - Age of index children who were subject to Case Management Review at time of index event were as follows: Under 1 year - 29% Between 1 year and 5 years - 17% Between 6 years and 10 years - 4% Between 11 years and 15 years - 33% 16 years and above - 17% n=24. • Sex - Gender of index child - Female - 54% Male - 46% (n=24).	analysis of CMRs with relevance to aspects of professional practice and ways of working in early help: **Early and sustained intervention (p47):** The study identifies the following points in relation to early and sustained intervention: - Few CMRs showed evidence of early recognition of significant risk factors, including issues such as adult mental health problems. - In other CMRs, primary care professionals either did not recognise risk factors or, where they had concerns, did not share these with children's social services or family members who could have kept the child safe - A piecemeal pattern of on/off engagement by social services and other agencies with families - Lack of appropriate information sharing by children's social services and other agencies - Lack of early intervention was particularly apparent in cases involving adolescents - Agencies were poor at addressing the impact of chronic neglect on children and intervening at an early stage - Failures to respond in a sustained way to extreme distress, manifested as risky behaviour, particularly among 'hard to help' adolescents. The report highlights 'The importance of early, more sustained and better coordinated intervention, not just for younger children but older children and adolescents' **Child neglect (p48):** No data relevant to early help specifically **Thresholds for intervention (p49):** No data relevant to early help specifically.	**of internal validity:** + **Overall assessment of external validity:** + Study contains information relevant to guideline. **Overall validity rating:**+ Overall, there is a lack of description of how thematic analysis was undertaken.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Ethnicity - All index children were White and had been born in Northern Ireland. • Religion/belief - Not reported. • Disability - Of sample of 24 children: - 8 had a mental health disability, 8 had an intellectual disability or ADHD, 6 participated in drug or solvent misuse, 5 participated in alcohol misuse, 3 had a physical illness, 1 had a physical disability, and 1 had a sensory impairment. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Types of abuse reported under 'indicators of concern' - Family history of child neglect identified in 7 cases - Family history of child emotional abuse identified in 2 cases, family history of child physical abuse identified in 6 cases, family history of child sexual abuse identified in 11 cases. • Looked after or adopted status - Looked after status: looked after at time of index event - 21%, previously looked after - 17%, never looked after - 62% n=24. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported.	**Communication and information sharing between professionals:** No data relevant to early help specifically. **Recording and record keeping:** No data relevant to early help specifically. **Compliance with established policies and practice**: No data relevant to early help specifically. **Assessment and analysis of information:** To be reported under NCCSC questions on assessment. **Supervision, staff support and training:** To be reported under NCCSC question on organisational factors. **Keeping the focus on the child:** No data relevant to early help specifically. **Organisational and staffing context:** To be reported under NCCSC question on organisational factors - Inter and Intra agency working.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Sample size: 24 Case Management Reviews analysed.		

8. Domian EW, Baggett KM, Carta JJ et al. (2010) Factors influencing mothers' abilities to engage in a comprehensive parenting intervention program. Public Health Nursing 27: 399–407

Research aims	**Study aim:** 'To identify possible factors influencing the ability of mothers perceived to be at the highest risk for child maltreatment to engage in a home visitation program' (p399). **Methodology:** Qualitative. **Country:** USA. **Source of funding:** Government - The National Institutes of Child Health and Development; Centers for the Prevention of Child Neglect, US.
PICO (population, intervention, comparison, outcomes)	**Participants:** • Data source was professionals/practitioners. Coach interventionists of parenting programmes. • Participants in the intervention were caregivers and families. Parents at risk for child abuse and neglect engaged in parenting programmes. **Sample characteristics:** The 4 coaches spoke about their experiences of working with the following sample of families: • Age - Mothers aged 15–35 years of age (6 teenagers); 3 coaches middle aged adults, 1 coach in late 20s. • Sex - mothers; all coaches female. • Ethnicity - Eight mothers African American, and 1 mixed African American and Caucasian. Coaches: 3 Caucasian and 1 African American. • Religion/belief - Not reported. • Disability - 4 mothers had learning disabilities, 6/9 mothers had low literacy levels. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - None of the mothers had completed high school. Mothers inadequate family and/or poor social supports.

	• Type of abuse - Mothers with a history of previous reports of child neglect or abuse with other children; problems with anger management; involvement with alcohol and/or street drugs; history of domestic violence and/or psychological trauma. • Looked after or adopted status - 1 teen mother lived with a foster family. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - not applicable. • Intervention number - not applicable. • Sample size – 4 coaches, reporting on work with nine families **Intervention:** The context for this study is that of 'My Baby and Me' multisite home visitation intervention study.
Findings.	**Narrative findings – qualitative and views and experiences:** The authors identified three themes which reflected coaches' perceptions of mother's ability to engage in programmes. **Theme 1:** 'Mothers struggle to meet the emotional needs of the self and the child' 1a. Difficulty in identifying emotional needs: Coaches in the study reported that mothers often had difficulty in recognising how their own emotional wellbeing impacted their child's emotional and behavioral state and these difficulties were often mirrored in the mother-child relationship. An unemployed and isolated 30-year-old mother with an infant daughter and a 2-year-old son said: '… if the two-year-old is acting out … and not really doing what she wants him to do, it's his fault, it's not because she's been yelling and kind of really harsh with him' (p402). 1b. Confusion in emotional caretaking. Coaches in the study found that some mothers were unaware and confused of her own and the child's emotional needs. 'They [mothers] expect their babies to understand their needs as women, as mothers, as adults … that their babies should be able to understand those things …' (p403). 1c. Difficulty in trusting others. Some high risk mothers were reported to be less able to trust and communicate openly with coaches, and ask for help. **Theme 2:** 'Mothers lack support in navigating complicated and stressful life events' 2a. Coaches reported multiple deprivations and support in a non-nurturing environment, leading to isolation that prevented them from managing stressful life events. 2b. A future of uncertainties. Coaches in the study reported that mothers lacked the skills or resources necessary to deal with complicated and stressful parent-child situations: ' … none of' [mothers] have vehicles …. faced with all these stressors …. they don't have any way to escape …. get out of the house or … go to the store …' (p403). Coaches also perceived that, due to instability in living conditions, mothers moved around a lot. 2c. Isolated in life struggles. Coaches reported a lack of financial and personal support for mothers which meant they do it all on their own. Additional psychological stressors such as sexual assault and abuse, depression, learning disabilities, minority status etc. could compound the situation.

	Theme 3: 'Mothers' consistency with program engagement is mediated through a trusting and caring relationship with coaches' 3a. Need for genuine caring. Coaches reported some mothers difficult to engage. Specific programme materials and interventions could either increase involvement or cause psychological conflict for the mothers, resulting in disruption to engage in parenting programmes. All the coaches felt that attentive human contact, providing opportunities for genuine communication with another person, and allowing mothers to express ideas were important to engage mothers in parenting sessions. '… just being able to relate to their situation and 'basically how they're feeling that day' increased the mother's involvement with program materials' (p404). 3b Need for a committed coach relationship This required coaches to meet the mother 'where she is' physically, emotionally, and mentally, and to make her feel valued during all interactions and interventions. Such relationships not easily accomplished due to emotional burden and time commitment. Coaches also felt that they could better support the mothers if the mothers could identify with the coach, for example on grounds of age, race, history of childbearing, or other experiences. This study showed that vulnerable mothers' levels of engagement can be mediated by professionals who are sensitive to the individualised factors influencing the mothers' lives and decisions. To develop a trusting and caring relationship, it is important that home visitor coaches/nurses understand the psychological and the contextual factors that influence mothers' ability to engage, and to achieve the continuous process of engagement by supporting mothers to explore and discover self-care strategies and ways to manage life struggles. Study also supports some well-known barriers to engagement, including fear of 'the system', and the difficulty in parenting in the face of extreme poverty, and lack of emotional and financial support. Limitations: Teen mothers and older mothers would have different life experiences and developmental needs Coaches' perceptions only, not mothers.
Overall validity rating.	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** + **Overall validity rating:** +

9. Easton C, Lamont L, Smith R et al (2013). 'We should have been helped from day one': A unique perspective from children, families and practitioners. Slough: National Foundation for Educational Research

Research aims	**Study aim:** To explore effective early interventions in response to neglect and families and practitioners views on these. The research question was: 'How do we effectively support families with different levels of need across the early intervention spectrum to engage with services within an overall framework of neglect?' (piv). **Methodology:** Qualitative study - Although the Local Authorities Research Consortium provides the overall research

	question, the local authorities carried out their own research. Three indicated that they used quantitative methods although it is not clear if this data is included in the report. **Country:** UK. Nine local authorities in England: Bracknell Forest Council, Coventry City Council, Hertfordshire County Council, Kent County Council, Portsmouth City Council, Solihull Council, Telford and Wrekin Council, Wolverhampton City Council and Warwickshire County Council. **Source of funding:** Not reported.
PICO (population, intervention, comparison, outcomes).	**Participants:** • Children and young people – Children and young people experiencing neglect. • Caregivers and families - Families in which child neglect was a problem. • Professionals/practitioners - Practitioners worked in a range of fields including '... education, health, early years settings and authority services ...' (pv) This included head teachers, school nurses, police officers, targeted youth support workers, domestic abuse workers, Home-Start workers, etc. **Sample characteristics:** • Age - Children and young people: '... aged up to 11 ...' and '... aged 12+ ...' (p 45). No details are provided regarding age of families and caregivers, or practitioners. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - The study focused on children experiencing neglect which met definitions used in Southampton's Local Safeguarding Children's Board 'Really Useful Guide to Recognising Neglect' (Southampton 2012): 'Level two, related to families where the parent/s mostly met the child's needs. Level three, where children had some unmet needs; lived in a family home that lacked routines; had parents with poor awareness of safety issues; and the child received limited interaction and affection. Level four, these were families in which adults' needs were put before the child's, and where the child had low nutrition and scarce stimulation' (p1). The report did not aim to '… consider cases where children were at significant risk of harm and should be being supported by statutory services' (p1). • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** 9 local authorities provided data from –

	• n=105 practitioners. • n=25 parents or carers. • n=15 children and young people.
Findings.	**Narrative findings – qualitative and views and experiences:** 'How do authorities support families experiencing neglect?' (p9). Practitioners reported a range of services which they felt were appropriate when responding to families experiencing neglect at level two. These included parenting programmes (e.g. Triple P) or child and parent support groups, children's centres, the Family Intervention Project, Home-Start, housing services (i.e. for help with tenancy issues), and mental health services. Gaps in provision - A number of practitioners working in education suggested that the need for a CAF to access services was problematic and that the need for parental consent also caused difficulties. Other practitioners reported that families had difficulties in accessing health services such as child and adult mental health services. Some early years practitioners suggested that new parents (at level 2) should receive greater assistance to cope with the '… emotional upheaval …' (authors, p11) of a newborn baby. Practitioners felt that there were also gaps in education support for families at levels three and four, such as a lack of family support workers in schools, or the small number of Emotional and Behavioural Difficulties placements. Gaps in parenting provision - Practitioners felt that parents in families at levels 2, 3 and 4 needed help with basic parenting skills and managing budgets. They also suggested that help needed to be offered at an earlier stage and attempt to address the issue of stigma (suggestions included a coffee morning). For families at levels 3 and 4, practitioners felt that support should be provided in the home in some instances, and that some 'chaotic' families were unlikely to engage with group based support. Practitioners also commented on the long waiting lists for parenting programmes. 'Perceived reasons for gaps in provision' (p12). Practitioners felt that other professionals were often unaware of the extent of support available for families at a local level. Other reasons given include the failure to identify family needs early enough, high thresholds, reluctance to enable early intervention and use the CAF process, and poor information sharing. Practitioners also emphasised the lack of preventative help which they felt would prevent some problems escalating to level 4. Suggested methods of addressing these gaps included providing information to families regarding available support, developing clear action plans that all practitioners are aware of, ensuring that workers assigned to families are consistent, delivering more support in the family home, improved training, enhanced school provision of early intervention (i.e. through family support workers), better coordination of multidisciplinary support (one practitioner noted that a family had received five visits from different professionals in one day. One practitioner noted that they supported parents to access services by accompanying them on their first visit to a service. The authors report that practitioners most authorities felt that gaps in service provision were not specific to neglect, but

	rather to '… early intervention support in general' (authors, p18). A number suggested that practitioners in children's and adult services needed to cooperate more closely and suggested a single directorate as the solution whilst others felt that some of the families they dealt with actually had very high needs which they lacked experience in dealing with. Although the authors report that the majority of families felt that there were no gaps in service provision and were happy with the support they had received, a small number from 2 local authorities identified issues such as gaps in health and education services (e.g. signposting to other services), the time-limited nature of support and a lack of information as particular problems. 'Do practitioners feel prepared to meet families' need?' The authors report that only a few practitioners were asked this question specifically but those who did identified issues such as their reliance on other professionals for information about families, a need for greater multi-agency working, inconsistent ways of working with families and the failure to initiate CAF processes at an earlier stage as challenges they faced. Practitioners were also asked whether there were any unmet training needs in relation to supporting families experiencing neglect. The authors report that '… a slightly larger proportion indicated that they or their colleagues had unmet needs' (p18). They also note that practitioners in four local authorities were specifically concerned with children's social care staff with some suggesting that social workers were too lenient with regards to poor behaviour which then escalated to more serious issues. 'How do practitioners think they can best meet families' needs?' (p19). Access – Practitioners felt that services and information about them needed to be more easily available to families and that universal services should play a greater part in preventative efforts. One practitioner felt that the need to make an appointment before accessing services was a problem and suggested that the use of 'open-door' policies. Multi-agency working – Practitioners felt that multi-agency working needed to be improved and suggested that better communication and information-sharing would help in this respect (with one practitioner commenting on the value of co-location). Some also felt that practitioners needed to develop a common language which would help both practitioners and families. At one local authority, it was reported that social housing landlords were invited to multi-agency meetings, and one practitioner felt that housing practitioners could signpost families to appropriate services if they became aware of particular problems in the family home. Practitioner skills – The authors report that the majority of practitioners felt that their skills and the relationships they developed with families were a fundamental component of the support they provided. They suggested a number of key principles including taking the time needed to develop relationships, being persistent, ensuring that they are consistent in the work they do with families, being honest and confident enough to state when a caregiver displays unacceptable behaviour, fostering trust, being aware of services available, being non-judgemental, motivating families and setting realistic goals, and being flexible and available. The authors quote a number of practitioners including one who commented on the importance of consistency.

	'Being consistent in that support, being firm with them. Doing what you said you were going to do, making sure it was followed through … so they knew that you were being honest' (practitioner, p20). A small number of practitioners commented on the importance of acceptance of the practitioner by families and the need for a good first impression: 'It's a bit like a job interview, you have got a minute to impress, if they don't like you, you're not going anywhere' (p21). Interventions and strategies which practitioners felt were useful included: parenting programmes (e.g. Triple P) and the 'Solihull approach' (courses designed to help parents understand their child's behaviour); Family Intervention projects; location of family support workers in schools; home visits; contacting families when they miss appointments; accompanying parents on their first visits to courses or services; and ensuring that support is consistent and that there are no breaks in service (e.g. when families fall below thresholds). Families were asked which aspects of support they had found most helpful. They identified emotional support, practical and financial support, programmes and clubs for children and young people and parenting support or courses as particularly helpful. Emotional support – The authors note that parents 'commonly' specified that emotional support had been the most valuable type of help they had received. They stated that having someone to listen to them who was non-judgemental had led to improvements in their parenting: 'Having someone here to support me emotionally and practically who did not judge me or my situation was great. Knowing I had support and could phone up at any time to ask for advice was great too. It gave me the strength to work at being a better parent' (parent, p22). Some parents also appreciated the fact that practitioners acted as an advocate and supported them in difficult meetings. Whilst others commented on the emotional support their child had received which had improved their behaviour. Practical and financial support – The authors note that many parents valued the practical and financial assistance they received such as help with repairs to their house or to obtain household goods and advice regarding finances. Programmes and clubs for children and young people – Some parents reported that clubs which their child had attended had had a positive effect on their behaviour. These included youth centre activities, anger management programmes, mentor support, etc. Children and young people also reported that these types of activities had been beneficial. Parenting programmes and support – The authors report that parents appreciated support which helped them to improve their parenting, which they reported had increased their confidence and self-control: 'They helped me feel confident that I could be a good parent and take care of my children' (parent, p23). The authors also note that a small number of children and young people reported that this type of support had led to improved parenting with their parents responding in a calmer manner and being able to solve problems themselves.

	Children and young people also appreciated having someone to talk to, which some reported was the most helpful type of support provided. 'Barriers and enablers to supporting families'/'Why families do not engage with services' (p24). The authors report a number of themes which came up when practitioners and families were asked to comment on barriers and facilitators to engagement. Misconceptions – Families misunderstood the role of social services and are afraid of their child being removed from the home. Families are also afraid that other professionals they come in to contact with will contact social services, as one professional stated: 'I know a woman who was scared to go to her GP about her low mood because she was afraid it would lead to a social worker taking her child away. So, there's that real misunderstanding of what social workers do' (practitioner, p25). Practitioners also stated that the stigma attached to receiving support from social services prevented parents from engaging. Families also commented on stigma and feared that asking for help would lead to the removal of their child, as one parent commented: 'I was petrified that if I asked for help my kids would be taken away from me […] and it's happened to my friends. I was really scared but also really desperate for that help' (parent, p25). Previous experience with services – Both practitioners and families noted that prior experience had an effect on the likelihood of families engaging with services in the future and a number of families commented on perceived attitudes of professionals which they had found unhelpful such as being judgemental or condescending. Practitioners also suggested that the tendency for other professionals to act as 'authority figures' can act as a barrier to engagement. 'Processes and resourcing' (p26) – Practitioners reported that time-limited services and the use of jargon were potential barriers to engaging families. Some families reported that thresholds, particularly in relation to children's social care and mental health services, prevented them from getting assistance. In relation to family level issues, practitioners suggested that parent's feelings of being judged was a barrier to engagement, which children and parents also suggested was a concern. Some also viewed professionals as unlikely to help or listen to them, particularly if they had unsuccessfully sought help in the past. Some parents were also reluctant to seek help through fear of being a burden. The authors asked parents what helped them to engage and they report that the majority stated that being aware of the help available to them was important. They suggested that services needed to be promoted and advertised to a greater extent and identified practitioners who supported and listened to them as another factor which would help. Other suggestions included an open evening to address misconceptions about social services and fear of stigmatisation, and support which was coordinated with one parent explaining: 'It's helpful to get everyone together because people can't cope with lots of different phone calls etc.' (parent, p28).

	When asked to discuss barriers to supporting families, practitioners reported difficulties associated with multidisciplinary working including, other practitioners who were unaware of available support and '… unresponsive to proactive family requests for help …' (authors, p29). Practitioners also reported that 'access' was an issue and specified that low numbers of home visits and the need for family consent in order for early interventions to commence were particular difficulties which they faced. Some also felt that the lack of alternative sources of support for families who do not meet thresholds can make re-engagement more challenging. When asked about the help they received, the authors report that the majority were positive and recognised that they had needed assistance. Some felt relieved that they were receiving help and others were glad that they were being 'listened to'. The authors also report that the majority of children and young people were happy about the help their families had received although some had initially been nervous about this. Parents were also asked whether they had received help at the right time and the authors state that: 'Around two-thirds of the parents stated that they would have liked the help sooner. Some parents recognised that this was partly because they were not aware of the organisations that could help. However, others felt that they were not being listened to or that services (particularly education and children's social care) were not acting quickly enough to help them' (authors, p31). The authors also report that the majority of families perceived that the support they had been given had led to positive changes in their circumstances such as a more stable home environment, improved child behaviour, and improved parental mental health. In contrast, there were some families who reported that the assistance had had only minimal impact, particularly if it had only recently been provided. The authors report that as parents stated that they would ask for help again from the practitioner or agency they were currently receiving support from this is an indicator of the value which it had. Other parents reported greater awareness of the support they could access and greater confidence in doing so. 'Families and the 'revolving door' (p33): The authors report that practitioners saw the 'revolving door' issue as partially a result of the way in which services were delivered: Time of case closure – Practitioners felt that support was withdrawn from families before their resilience had developed sufficiently and some suggested that the withdrawal of support should be tapered rather than immediate which would allow practitioners to monitor progress. The authors note that both families and practitioners felt that services should not be time-limited. Communication between services – Practitioners identified poor dialogue between professionals as an issue, with some working in education noting that they were not informed when an intervention had ended. Working with families: Some practitioners believed that support too often focused on the 'symptoms' of neglect rather than the underlying causes and the authors quote a practitioner comment as 'typical' which stated that: 'You will not break the cycle just by addressing the symptoms' (practitioner, p34). The authors suggest that practitioners need to work with families using strategies which enable them to build capacity to support themselves and change the underlying behaviours which contribute to the cycle of neglect. Practitioners also suggested strategies such as working with

	the whole family (e.g. addressing adult mental health problems), developing clearer plans for change with appropriate outcomes, encouraging parents to think about their own behaviour rather than focusing on 'problem' behaviours of their child, etc. Practitioner views on engaging extended families were mixed, with some who felt that this was a helpful way of addressing family issues whilst others felt that this tactic may in some instances be harmful. Some practitioners felt that supporting families to engage with their community was an appropriate way of extending their support network, however this was contrasted with responses from other practitioners who thought this was inappropriate and should not be seen as a substitute for necessary services.
Overall validity rating.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall validity rating:** - The failure to provide details regarding the analysis process and the characteristics of the sample are significant limitations.

10. Fernandez E (2004) Effective interventions to promote child and family wellness: A study of outcomes of intervention through Children's Family Centres. Child and Family Social Work 9: 91–104

Research aims	**Study aim:** To '... to evaluate the impact of family support interventions by comparing the views of families and their caseworkers with respect to the perceived benefits and outcomes of the interventions in the context of changes in family functioning and parent–child relationships, and the extent to which changes led to reduced involvement in protective services' (p91). Data extraction for this study has focused on the perceived benefits of the intervention (qualitative data from families and caseworkers), as most relevant to review question 14. **Methodology:** Part of a mixed methods study. Only qualitative data (from semi-structured interviews) have been extracted and the study has been appraised as a qualitative study. **Country:** Australia - Sydney.
PICO (population, intervention, comparison, outcomes)	**Participants:** • Caregivers and families - Families identified as being at risk for child abuse and neglect and were in receipt of services from Children's Family Centres. The participant families had been referred for a range of reasons. 28% had been referred (including some self-referrals) primarily for assistance with accommodation and a number of families had experience periods of homelessness. 24% were referred primarily because of difficulties managing their children's behaviour. 28% had been referred due to child protection concerns or wished to place their child in temporary

	foster care. 21% had been referred due to '... relationship issues ...' (p93). The authors note that in all of the 29 families the women had experienced domestic violence and '... needed assistance living with the aftermath of the violence' (p93). The sample also included children whom statutory workers had assessed as having been abused. • Professionals/practitioners - Caseworkers delivering services through Children's Family Centres to families identified as being at risk for child abuse and neglect. **Sample characteristics:** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Data is not provided but the authors note that the sample included parents with learning disabilities. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - The authors note that a '... significant number of the families were characterized by sole parenthood, social isolation, homelessness, debt, and alienation from family networks' (p93). • Type of abuse - Not reported specifically: '... children assessed by statutory workers to have been abused or at risk of abuse and of entering care, or of being restored from care ...' (p 93). • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** n=29 families were chosen who had been referred to the service, however 3 families then declined to participate as a result of crises they experienced at the time. **Intervention:** Children's Family Centres are run by Barnardo's Australia and provide integrated family support. The services they provide are intended to be holistic, non-stigmatising and to encourage families to proactively seek assistance. The aim of the programme is to reduce risk factors and enhance protective factors. Specific services include home visiting, day care, semi-supported accommodation, counselling, group sessions, respite care and crisis services. The centres also run Temporary Family Care - a crisis service open on a 24 hour basis which can provide crisis responses and respite care.
Findings.	**Narrative findings – qualitative and views and experiences:** The authors note that caseworkers identified positive changes in the families they worked with, for example in parent's attitudes and decision-making. The authors also report that the '... general picture ...' from interviews with families was that the services they received were useful in both emotional and practical terms. The authors note that parents described approaches which they had found helpful such as '... listening, being non-

	judgemental, respectful and accessible ...' (authors, p 100) They quote one parent who stated that: 'Well they've helped us a lot. They've given us "respect", like normal people just run you down – don't give a damn how you feel and that, and they've just given us that support, just being able to cope, and that … just being able to talk – that helps' (parent, p100). The authors also report that families' positive views on tangible benefits such as monetary assistance or help with housing was often followed by comments on the less concrete benefits associated with the programme such as the development of support networks, or the potential to make friends: 'Yeah, they've given us ideas - how to change things and just - be more relaxed and that, they've helped out financially with the power bill, and food wise – and stuff and just introducing me to um the Mother's Group – that really helped cause I've made a really good friend out of it, and we see each other all the time' (parent, p100). The authors state that '... sympathetic and accessible professionals were important to many parents ...' (p100) and that the ability to share problems with other parents and discuss what was 'normal' and what was a 'problem' were also valued. Parents also reported that they had become closer to their children and that their parenting had improved. The authors asked parents what they liked and disliked about the service and they report that many commented on the accessibility of the service, with one parent commenting: 'Just her coming out and um just having someone that'll come out to your house cause there's time there - no one used to come to the house, you know, it was like just me and him, all the time with the kids and it just got monotonous. Just having a person come into your house and respect ya, and everything - that even helped and not criticize ya' (parent, p100). Parents also felt that they had benefitted from the intervention and the authors highlight such views as 'typical': 'I have to say, at the end when they played a pro-active role in trying to help was really good, rather than waiting to have me call and scream for help – it was good when they were more – jumping on things before things got bad – that was good. So, it was good to have her phoning and saying "how are things going" and that, you know, rather than to have me phone and ask for help' (parent, p100). 'They have taught us all how to live with each other and just how to cope with all the little things that come up in everyday life. There hasn't been any drawbacks. I don't know where I'd be without them, I know for a fact that I wouldn't have my kids, so …' (parent, p 101). Parents also welcomed support which was tailored to their individual needs and supported their parenting goals. They also appreciated support which was delivered in their home and combined tangible assistance with emotional support and education. Parents also valued the role which the service played in mitigating their lack of family and social networks. In addition

	to positive comments regarding the service, the authors report that workers and parents reported ongoing issues of concern. Some workers felt that it was important to remember that the positive changes which families had made had taken place in a short timescale, and that the problems they faced were 'intractable' (authors, p101). Family support workers still had concerns regarding the use of physical discipline and issues in relation to attachment and the authors state that transcripts of interviews with workers show the difficulty in supporting families whilst also remaining aware of child protection issues and the point at which a threshold has been reached.
Overall validity rating.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall validity rating:** - The failure to include any detail regarding the analysis process and only minimal details on the characteristics of participants and settings are a key limitation.

11. Girvin H, DePanfilis D, Daining C (2007) Predicting program completion among families enrolled in a child neglect preventive intervention. Research on Social Work Practice 17: 674–85

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To build a model to predict which families complete the Family Connections programme, a targeted preventive intervention delivered to families who meet child neglect criteria. The model included consideration of three variables relevant to aspects of professional practice and ways of working: -	**Participants:** Caregivers and families - Families who meet child neglect risk criteria with at least one child between the ages of 5 and 11. There were 4 eligibility criteria - 1. Concern from a 'referring person' that at least one of 19 neglect subtypes was present at a level below that required for CPS investigation. Subtypes included unsafe housing, poor health care, etc. 2. Concern that at least 2 other risk factors related to the child or caregiver were present. These included behaviour problems; physical, learning, or developmental disabilities;	**Quantitative data - Satisfaction with services -** **Results from bivariate analysis:** • Intervention group (3 or 9 month programme): Significant difference in completion, in favour of 3-month group (3 month group: 61 competed, 3 did not; 9 month group 51 completed, 19 did not), $\chi 2$ [1, n=136] = 10.745, p=0.001. • Helping Relationship Inventory-Client - Interpersonal subscale: (Independent sample t tests. Significance assessed after Bonferroni correction. Only p values of less than .001 are significant). Caregivers who completed the programme reported higher scores on the HRI-C interpersonal score than those who did not complete the programme: Programme	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:**+

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
• Length of intervention. • The helping relationship between worker and client. • Satisfaction with interaction with worker. Data have been extracted in relation to these variables only. Other variables considered related to characteristics of the participants or their families, rather than aspects of professional practice. **Methodology:** Other - Correlational design measuring the association between a range of independent variables and service completion - and indicator of satisfaction with services. The study has been critically appraised using a tool designed for cross-sectional studies, which are closely related to correlational	more than three children; unemployment; mental health problems; domestic violence; homelessness; etc. 3. Not currently involved with CPS. 4. Caregiver willingness to participate. **Sample characteristics:** • Age - Age of primary caregivers ranged between 19 and 72 years (M=36.88, SD=12.21). Age of children ranged from newborn to 20 years old (M=8.34, SD=4.05). • Sex - The majority of caregivers were female = 97.8%. • Ethnicity – The majority of caregivers were African American = 87.5%. The majority of children were African American = 86.4% • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Never married = 67.6%, unemployed = 59.6%, mean educational level = 10.80 years (SD=2.27), mean income level = $9,931.83 (SD = $5,926.18), number of children per family ranged from 1 to 9 (M=3.14, SD=1.63). Children living with their mothers=78.0%. • Type of abuse - The most frequent concerns (in relation to neglect) at	completers M=39.18, SD = 8.89; Caregivers who did not complete programme M=28.87, SD=11.71; t=−4.788, p<.0005. The size of the effect was large (Cohen's d = 1.11). • Helping Relationship Inventory-Client - Structural subscale: (Independent sample t tests. Significance assessed after Bonferroni correction. Only p values of less than .001 are significant). Caregivers who completed the programme also reported higher scores on the HRI-C structural component. This was not significant according to the Bonferroni-corrected criterion, but did show a large effect size: (Programme completers M=40.19, SD=8.40; Caregivers who did not complete programme M=32.13, SD=12.20; t=–3.026, p=.005, Cohen's d=.89). • 'Satisfaction with interaction of workers' as measured by Parent Outcome Interview: (Independent sample t tests. Significance assessed after Bonferroni correction. Only p values of less than .001 are significant). Caregivers who completed the program reported significantly higher satisfaction with workers, with a large effect size: (Programme completers M=2.44, SD=0.62; Caregivers who did not complete programme M=1.82, SD=0.75; t=–4.213, p< .0005, Cohen's d=.97). **Results from multivariate analysis (logistic regression):** NB Only data from the final model (block 3) has been extracted as this model contains all variables of interest to NCCSC question 14. The significance criterion used was p<0.005 (note this is not the usual significance level, but is stated in paper on p.681).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
designs. **Country:** USA - Baltimore. **Source of funding:** Not reported.	intake were - delay in accessing children's mental health care (35%), unstable living conditions (24%), inadequate supervision (23%), the majority of families also had employment problems (68%) or child behavior problems (66%). • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** This study represents part of a larger randomised controlled trial which aimed to evaluate the effects of the Family Connections programme. This used an experimental design to randomly assign 154 families into four conditions: 1. Family Connections programme for 3 months. 2. Family Connections programme for 3 months with a group intervention. 3. Family Connections programme for 9 months. 4. Family Connections programme for 9 months with a group intervention. 'Because of poor compliance with the group intervention, this article reports only on a comparison of 70 families who were assigned to FC intervention for 3 months versus 84 families assigned to receive FC intervention for	• Intervention group (three or nine months programme): This was a significant predictor of completion within the model (B=1.995, p=0.003) • Helping Relationship Inventory-Client - Interpersonal subscale: Scores on the helping relationship inventory interpersonal subscale were also significantly predictive of completion (B=0.085, p=0.049). • 'Satisfaction with interaction of workers': This was not a significant predictor of completion (B=0.650, p=0.295). • Helping Relationship Inventory-Client - Structural subscale – Not reported. **Narrative findings:** **Results from bivariate analysis:** • Intervention group (3 or 9 month programme): There was a significant difference between those who completed the programme and those who didn't in relation to treatment group assignment, with families assigned to the three month programme being more likely to complete than those in the none month group. • There was a significant difference in scores on the Helping Relationship Inventory-Client - interpersonal subscale - with those who completed the programme reporting higher scores than those who did not. • There were no significant differences in scores on the Helping Relationship Inventory-Client - structural component subscale although there was a large effect size with caregivers who completed the programme scoring higher than those who did not. • There was a significant difference between scores	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	9 months, combining those with and without the group intervention' (p677). n=70 received treatment for 3 months (some of whom received the enhanced version of the programme which included a group intervention). n=84 received treatment for 9 months (some of whom received the enhanced version of the programme which included a group intervention). n=136 families completed interviews at service termination. **Intervention category:** Multi-component intervention. **Intervention:** Family Connections is a targeted preventive intervention aiming to reduce the risk of child neglect. It uses a '... family-centered model of practice ...' (p677) and is based on nine principles such as the helping alliance, individualized family assessments, outcome-driven service plans, etc. It is delivered by social workers and social work interns '... in the context of their neighborhoods ...' (p678) using a manual. The main components of Family Connections are; home-based family interventions (e.g. assessments and	on section 11 of the Parent Outcome Interview (satisfaction) with caregivers who completed the programme reporting higher scores. **Results from multivariate analysis (logistic regression) - final model (block 3):** • Treatment group and scores on the Helping Relationship Inventory-Client - interpersonal subscale were significant predictors of programme completion. Families assigned to the 3 month intervention were 7.35 times more likely than those assigned to the 9 month intervention. Each 1 point increase on the Helping Relationship Inventory-Client - interpersonal subscale increased the odds of completion by 1.09. • Client satisfaction measured using section 11 of the Parent Outcome Interview was not a significant predictor of programme completion. The authors conclude that '… the FC intervention emphasises the formation of helping alliances between the social worker or social work intern and all family members … Considerable effort is made by the program to teach and model methods for forming helping alliances between social work interns and family members. The findings of this study suggest that those efforts are important and should be continued as part of the implementation of intervention' (p683). The authors further conclude that the study suggests that families may find it easier to complete services designed for a shorter interval.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	service planning, counselling, etc.), coordination of services and referrals to services to address risk factors such as substance abuse, and to enhance protective factors, e.g. through mentoring programmes) and 'multi-family support recreational activities'. n=154 families were randomly assigned to one of four treatment conditions: 1. Family Connections programme for 3 months. 2. Family Connections programme for 3 months with a group intervention. 3. Family Connections programme for 9 months. 4. Family Connections programme for 9 months with a group intervention. **Outcomes measured:** Satisfaction with services - The authors aimed to build a predictive model of service completion based on variables noted in the literature. The dependent variable was service completion and the independent variables were: caregiver age, intervention group, Child Protective Services status, history of drug use, depressive symptoms, everyday stressors, 'the helping relationship', and 'satisfaction with interaction of workers'. As the NCCSC review question focuses on professional		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	practice or ways of working only findings relating to the following variables have been extracted: • Intervention group (3 or 9 month programme). • 'The helping relationship'. Measured using the Helping Relationship Inventory-Client (Poulin and Young 1997) at termination of programme. This is comprised of two subscales; the interpersonal component which measures interpersonal connections or bonds; and the structural component on which higher scores indicate collaboration and '... clarity about the purpose of the worker–client relationship' (p679). • 'Satisfaction with interaction of workers' measured using Section 11 of the Parent Outcome Interview (Magura and Moses 1986). Self-reported at service termination.		

12. Krysik J, LeCroy CW, Ashford JB (2008) Participants' perceptions of healthy families: A home visitation program to prevent child abuse and neglect. Children and Youth Services Review 30: 45–61

Research aims	**Study aim:** To explore the views of participants currently enrolled in a home visitation programme (Healthy Families Arizona) with specific regard to their '... perceptions of the intake process, the program's purpose, and the relationship between the home visitor and the participant'. **Methodology:** Qualitative. **Country:** USA.

	Source of funding: Government.
PICO (population, intervention, comparison, outcomes)	**Participants:** Caregivers and families. Participants of the Healthy Families Arizona programme. **Sample characteristics:** • Age – Administrators of the programme requested that identifying information such as age and marital status were not collected. ('Owing to the random selection of participants from each site, it is expected that the sample reflects the population of those receiving Healthy Families services', p 48). • Sex - Not reported. • Ethnicity - Approximately 54% of participants were Hispanic; 22% were white, non-Hispanic; 8% American Indian; 7% African American; 8% mixed-race identity, and 1% other. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position – 38% of the sample were teen mothers; 71% were not married upon entry to the programme, and 63% had less than a high school education. The median, gross annual family income was $9600. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - not applicable. • Intervention number - not applicable. • Sample size – 46 randomly chosen, currently enrolled families from 3 sites of the home visitor programme: 12 from the large urban site, 16 from the medium-sized urban site, and 18 from the site serving rural participants. **Intervention:** The context for this study is that of the Healthy Families America The 3 overarching goals of the Healthy Families America programme are: (a) to promote positive parenting, (b) to enhance child health and development, and (c) to prevent child abuse and neglect.
Findings.	**Narrative findings – qualitative and views and experiences:** Four questions were posed regarding the participants' experiences with the home visitation programme. **1. Their experience with the intake process** Immediate positive reaction: Study found that there was an immediate

	positive reaction in 18/46 or 39% of the sample. 'I felt wonderful. I wanted something because I knew that I was at the end of my rope. I drank through my pregnancy. I didn't need a baby. I was grateful for such a thing' (p49). Study found initial concern in 8/46 or 17% of the sample: 'Just the way that it was explained to me, I felt it was a child abusing thing. She wanted to come into the household to see if the baby was being fed and well loved. I felt like they thought I was going to abuse my child' (p50). Ambivalent - 3/46 or 7% of the sample: (p50). Neutral - 17/46 or 37% of the sample. **2. What they perceived to be the programme's primary** purpose Perceptions of the programme's primary purpose: Most participants said that the programme was 'to help, support, or provide services'. Some described the programme as being principally for the care of the child. 'It gives you good advice on what to do if your child needs to gain weight or what to do with discipline' (p51). Some viewed the programme's purpose as providing services for a specific target population. 'The program was designed to see how I am as a mother and how I am coping with the baby's behavior as a new mom.' 17% of the sample said that the programme turned out differently than expected. 'Well, when I got into the program I didn't know I would come to care for the Healthy Families person as much as I have. That is the plus in it all. I didn't realize that she would become such a big part of the family.' This suggested that the participants understand the purpose of the programme. **3. Their perceptions of the home visitor** 29/46 or 63% of the sample described their relationship with the home visitor as being more like a friend than a parent or teacher. 'She is real friendly and real polite. She'll ask me if I need anything and I think she really cares about me. She is a really nice person. She is more like a friend. She is not really like an authority' (p54). When participants were asked about what they liked about the home visitor, the authors report that three fundamental themes emerged: 1) Factors attributed to personal qualities of the home visitor, e.g. being a 'caring person' 2) The forms of concrete help they provided, e.g. help with taking children to appointments 3) Appreciation of home visitation component. What participants liked least about their home visitor (4/46 of the sample) 'Just that she needs to walk a mile in my shoes. To understand why I deal with my child the way I do. I have to spank him but it is something that has to be done. Every kid has to be spanked in my eyes. She thinks that is wrong. I don't see what is wrong with it' (p55). Almost all of the participants reported not feeling criticised by the home visitor. Most of the participants reported that the home visitors were not judgmental or critical in relating with them about their parenting abilities. **4. How their involvement in the programme changed over time** The study found that involvement: - remained the same for 21/44 or 48%. - increased (15 or 34%) (the quality of the participant's relationship with the home visitor appeared to be an important consideration in this increased levels of commitment) - decreased (8 or 18%) mainly due to changes in their life circumstances or they no longer needed the programme. 'Not as much as I was before. It is kind of like they come to me every two weeks now and they've done so much for me in the beginning, I'm getting by on my own now' (p58). Summary: Most participants felt that the understand the purpose of the programme to be 'to help, support, or provide services', some viewed it as being fundamentally for the care of the child. They described their relationship with the

	home visitor as being more like a friend than a parent or teacher, and that they had very close relationships with their home visitors which supported the intervention. They appreciated the specific personal quality of the home visitors as 'caring', 'listening' and providing concrete help when needed. This suggested that the participant and home visitor relationship was a central feature of the programme and this positive relationship would improve the delivery of home visitation services, strengthen the commitments and attachment of the families to stay with the programme, thus facilitating positive changes to address major risk factors for child maltreatment.
Overall validity rating.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** +

13. LeCroy C W and Whitaker K (2005) Improving the Quality of Home Visitation: An Exploratory Study of Difficult Situations. Child Abuse and Neglect 29: 1003–13

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: '... to use an ecological assessment model to obtain a better understanding of difficult situations that home visitors confront when implementing home visitation services' (p1003). The study was considered to be relevant to NCCSC question 14 as the findings provide information on areas of professional practice which need to be present or developed in order to work with difficult situations. The	**Participants:** Professionals/practitioners - Home visitors who deliver Healthy Families Arizona - a programme of home visiting provided to at risk parents (usually first time parents) based on the Healthy Families America model. Risk factors include poverty, a single parent, history of abuse, etc. Most home visitors had '... at least some college education (38.8% some and 45.9% college degrees)' (p1006). The authors report that most of those who had Bachelor's degrees had obtained them in '... human service fields such as social work, family studies, and psychology ...' (p1006) although no data in relation to this is presented. The authors also describe the home visitors as '...	**Quantitative data –** The authors used focus groups to create the Difficult Situations Inventory on which the survey was based. 'Included were specific situations in which individuals must respond effectively to be considered 'competent' ... Furthermore, these situations need to be 'problematical' to the degree that how to respond is not immediately apparent' (p1005). Top 15 most difficult situations identified by home visitors - mean ratings out of 5 (M) and standard deviation (SD). • 'Limited resources to help parents' - M 3.58, SD 1.14. 'Helping parents who threaten to commit suicide' - M 3.34, SD 1.23. • 'One person in the home is under the influence of alcohol or drugs' - M 3.34, SD 1.33. • 'Working in the homes during the summer heat' - M 3.31, SD 1.26.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall validity rating:** - The failure to determine the reliability and validity of the Difficult Situations Inventory is a significant limitation of this study.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
study aimed to reveal situations that home visitors find most difficult and, using factor analysis, group these in to themes. **Methodology:** Survey - A survey of 91 Healthy Families home visitors, based on an inventory developed through focus group research. **Country:** USA - Arizona. **Source of funding:** Not reported.	fairly experienced, having done home visiting for an average of 3.8 years; 70.4% were themselves mothers' (p1006). **Sample characteristics:** • Age - Mean 35.4 years, SD 10.4. • Sex - All participants were female. • Ethnicity - Caucasian 42.9%, Hispanic 28.6%, African American 4.4%, Asian American 1.1%, Native American 6.6%, mixed race 13.2%, other 4.4%. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** n=91 (participants who completed and returned the survey, representing a 90% response rate). **Intervention category:** Home visiting.	• 'When someone reports having given drugs or alcohol to children' - M 3.31, SD 1.26. • 'Responding to threats or dangerous behavior directed at home visitor' - M 3.19, SD 1.46. • 'Working with uncommitted families' - M 3.09, SD 1.10. 'Working with families that aren't motivated' - M 3.08, SD 1.13. • 'Dealing with family members who show up under the influence' - M 3.00, SD1.27. • 'Inability to contact parents' - M 2.99, SD 1.13. • 'Helping parents to change their parenting style' - M 2.98, SD 1.13. • 'Family members who are not motivated because of alcohol or drugs' - M 2.96, SD 1.12. • 'Families who are in constant crisis' - M 2.92, SD 1.00. 'Proving services in unsafe homes' - M 2.89, SD 1.23. 'Addressing domestic violence' - M 2.87, SD 1.08. Top 15 most frequent difficult situations identified by home visitors: • 'Working in homes during the summer heat' - M 3.97, SD 1.26. • 'Working with limited resources to help parents' - M 3.52, SD 1.21. • 'Working with teenage mothers' - M 3.48, SD 1.11. • 'Trying to create a confidential environment' - M 3.22, SD 1.42. • 'Knowing what activities to do during a home visit' - M 3.22, SD 1.6. • 'Working with parents whose decisions you don't agree with' - M 3.19, SD 1.18. • 'Working with families that aren't motivated- M 3.19,	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Intervention: Healthy Families Arizona - a programme of home visiting provided to at risk first time parents based on the Healthy Families America model.	SD 1.18. • 'Working with parent's emotional feelings (like sadness)' - M 3.18, SD 1.10. • 'Helping families when they are experiencing a crisis' - M 3.08, SD 1.08. • 'Working with uncommitted family members' - M 3.07, SD 1.23. • 'Working with parents who have different values' - M 3.04, SD 1.28. • 'Working with immature clients' - M 3.04, SD 1.16. 'Working with parents who are in denial about their problems' - M 2.98, SD 1.24. • 'Trying to collaborate with other agencies' - M 2.98, SD 1.24. • 'Inability to contact clients to set appointments' - M 2.98, SD 1.30. Home visitors who have worked with families experiencing domestic violence, substance abuse or mental illness: • Domestic violence - in the last year = 81.8%, M 5.09, SD 6.2. Domestic violence - in the last 30 days = 64.6%, M1.86, SD 2.5. • Substance abuse - in the last year = 82.7%, M 5.67, SD 8.6. Substance abuse - in the last 30 days = 67.5%, M 2.21, SD 3.1. • Mental illness - in the last year = 86.7%, M 4.84, SD 5.2. Mental illness - in the last 30 days = 78.5%, M 2.70, SD 3.1. Factor analysis of the difficult situations identified by the home visitors revealed the following five factors:	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		lack of clinical skill, family difficulties, parenting difficulties, personal difficulties, lack of experience. The items and their respective factor loadings are shown below: • Factor 1: 'Lack of clinical skill' – 'Working with a family member when they are under the influence' .79 'Dealing with a family member who shows up under the influence' .77 'Working with families when you know there is alcohol or drug use in the home' .73 'Working with families when someone reports giving drug or alcohol to children' .72 'Working with parents who deny alcohol or drug use' .68 'Providing services in unsafe homes' .64 'Inability to contact clients to set appointments' .62 'Working with parents not to change their parenting style' .62 'Working with parents who are not motivated due to alcohol or drug problems' .60 'Working with clients who are forced to receive services' .60 'Working with uncommitted clients' .59 'Not knowing how to intervene when parents use physical punishment' .48 'Dealing with polices or procedures that inhibit your progress' .45 'Working with parents when they have disclosed alcohol or drug use' .40 • Factor 2: 'Addressing family difficulties' – 'Knowing how to respond to child abuse with a family' .72 'Addressing domestic violence with a family' .71 'Knowing how to respond to child neglect with a family' .70 'Working with multiple social problems within a family' .70 'Knowing when to report a family to CPS' .68 'Knowing how to respond to domestic violence' .63 'Knowing when to report a family to	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		your supervisor' .54 'Using confrontation with families' .51 'Helping parents who threaten to commit suicide' .46 • Factor 3: 'Addressing parenting difficulties' – 'Working with families with limited understanding due to cognitive difficulties' .77 'Educating parents with mental health problems' .74 'Working with parents who have emotional feelings' .67 'Helping parents accept children 'the way they are' .65 'Working with limited resources to help parents' .61 'Finding strengths in families that you can use' .46 'Working with extended family members' .42 • Factor 4: 'Personal difficulties' – 'Dealing with personal frustration and failed efforts to help' .74 'Working with parents who have different values from your own' .72 'Working with parents regarding their sexual orientation' .65 'Trying to collaborate with other agencies' .64 'Making a successful referral for additional services' .60 'Working with parents whose decisions you don't agree with' .57 'Not understanding cultural differences' .58 'Feeling uncomfortable with the required paperwork' .54 • Factor 5: 'Lack of experience' – 'Knowing what activities to do in a home visit' .60 'Knowing how to intervene when problems arise' .58 'Not having enough experience to help parents' .53 'Not having enough experience to address mental health problems' .50 **Narrative findings:** The top 15 most difficult situations identified by home visitors were (in order of frequency) were: 'limited resources to help parents', 'helping parents who	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		threaten to commit suicide', 'one person in the home is under the influence of alcohol or drugs', 'working in the homes during the summer heat', 'when someone reports having given drugs or alcohol to children', 'responding to threats or dangerous behavior directed at home visitor', 'working with uncommitted families', 'working with families that aren't motivated', 'dealing with family members who show up under the influence', 'inability to contact parents', 'helping parents to change their parenting style', 'family members who are not motivated because of alcohol or drugs', 'families who are in constant crisis', 'providing services in unsafe homes', 'addressing domestic violence'. The top 15 most frequently occurring difficult situations which home visitors identified were: 'working in homes during the summer heat', 'working with limited resources to help parents', 'working with teenage mothers', 'trying to create a confidential environment', 'knowing what activities to do during a home visit', 'working with parent's whose decisions you don't agree with', 'working with families that aren't motivated', 'working with parent's emotional feelings (like sadness)', 'helping families when they are experiencing a crisis', 'working with uncommitted family members', 'working with parents who have different values', 'working with immature clients', 'working with parents who are in denial about their problems', 'trying to collaborate with other agencies', 'inability to contact clients to set appointments'. In the past year, over 80% of home visitors had worked with families where domestic violence, substance abuse and mental illness had been an issue.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Over 60% had worked with a family experiencing at least one of these issues in the last 30 days. The authors report that they identified through analysis (of 18 different factors) 5 five factors which explained 56.5% of the variance. The five factors were: • Factor 1: 'lack of clinical skill' (including working with clients whose participation is mandatory, being unsure how to respond when parents use physical discipline, being subject to policies and procedures which '... inhibit your progress ...' (p1010). • Factor 2: 'addressing family difficulties' (including 'confronting' families, working with families where there are a number of social problems, etc. Factor 3: 'addressing parenting difficulties' (including a lack of resources, identifying families' strengths on which to build, etc.) • Factor 4: 'personal difficulties' (including working with parents whose values are different, collaborating with other agencies, making 'successful' referrals, being unsure about cultural issues, etc.) • Factor 5: 'lack of experience' (including deciding which activities to use in home visits, deciding how to intervene when necessary, etc.) The authors conclude that 'many of these results suggest that home visitors may be overwhelmed by some of the complex situations they face' (p1009). They suggest that the five areas identified in the factor analysis could be thought of as core areas for training and supervision.	

14. Martin C, Marryat L, Miller M et al. (2011) The Evaluation of the Family Nurse Partnership Programme in Scotland: Phase 1 Report – Intake and Early Pregnancy. Edinburgh: Scottish Government

Research aims	**Study aim:** 'The overall aim of the evaluation is to assess the implementation of the FNP programme in Edinburgh and to use the learning from this to assess whether the programme can be implemented in other areas of the country. The evaluation focuses on three broad questions: - Is the programme being implemented as intended? If not, why not? - How does the programme work in Scotland (Lothian)? How do Nurses, clients and the wider services respond to the programme? What are the implications for future nursing practice? What factors support or inhibit the delivery of the programme? - What is the potential for FNP to impact on short, medium and long term outcomes relevant to Scotland?' (p20). It was considered that the research question on 'which factors support or inhibit the delivery of the programme' was relevant to this review question. We have extracted only the data based on the qualitative research conducted as part of the evaluation, as these were most closely aligned to our review question. **Methodology:** Qualitative study. Study overall uses mixed methods, but for the purpose of this review we have extracted only data obtained using qualitative research with clients and family nurses. This comprised 'in-depth interviews' 15 clients, a nominated significant other, and their family nurses. **Country:** United Kingdom. **Source of funding:** Government.
PICO (population, intervention, comparison, outcomes).	**Participants:** • Caregivers and families - Interviews conducted with 15 of the 148 women initially recruited to FNP. These were mothers aged 19 years or less at the point of conception. • Professionals/practitioners - Interviews were also held with the family nurses of the 15 clients interviewed. It is not reported how many individuals this comprised. **Sample characteristics:** • Age - Not reported, although all clients were under the age of 19 at time of conception and the majority were recruited by the 28th week of pregnancy. • Sex - All clients were mothers. Sex of family nurses not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported.

	• Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Clients n=15. • Family nurses – unclear. **Intervention:** The Family Nurse Partnership Programme is a prevent programme. Its goals are to: '… improve pregnancy outcomes, the health and well-being of vulnerable first time parents and their children, child development and families' economic self-sufficiency' (p13). The programme comprises an intensive, nurse-led home visiting programme, beginning during pregnancy and continuing until the child is two. Areas of support include: - preventative health practices - providing responsible and competent care - positive parenting - planning for the future. Basis of the programme is a 'therapeutic relationships' or 'alliance' (p14) between the family nurse and client. Theories underpinning the programme: - ecological theory - attachment theory - self-efficacy theory. The study describes the FNP 'core model elements', which are: - the visiting regime (frequency of visits is closely specified) - staffing requirements (professional and personal characteristics of family nurse) - client eligibility - the support organisation structures and processes required. The study also describes a number of fidelity 'stretch goals' which can help to maximise effectiveness. These relate to: - retention - visit 'dosage' (numbers and length) - coverage of different topics during visits (p15).
Findings.	**Narrative findings – qualitative and views and experiences:** Data relevant to aspects of professional practice in early help. Section 4 - Engaging and enrolling clients Section 4.3 Clients' and the delivery team's experiences of the recruitment phase (p27). **4.3.1 Finding out about FNP**: The study reports that the majority of clients were referred to the programme either by their midwife, or were contacted directly by a family nurse. **4.3.2 The engagement process:** The study found that the majority of clients who were approached decided to take part in the programme, but that clients did not always decide to take part immediately 'the process of engagement was lengthy for some clients' (p28). The study found that 'Nurses described the importance of taking time to include the mother or partner in the engagement process to answer any questions they may have had' (p28). **4.3.3 Deciding to sign up to FNP:** The study found that while some participants signed up with FNP because it 'sounded good', others had more clearly defined reasons for signing up, including wanting practical and emotional support, and having someone to talk to in confidence who was outside their family. The study also found that clients were keen to obtain information about parenting. One client said 'I didn't have a clue about anything' (p29). Concerns expressed included: - Worries that the family nurse 'sounded like a social worker' - Information being passed to other agencies - Fear of having baby removed. Client expectations of the service included: - Provision of information, advice

	and support about 'physical aspects of pregnancy and childbirth, ante-natal care and ante-natal classes' (p29). - Advice and support after birth - Information about role as parent - How to take care of a baby - Improve confidence - Provision of help with college applications or employment. The study reports that family nurses emphasised the importance of giving potential clients time to make their decisions about whether to enrol. **4.3.4 Recruitment schedule:** No relevant information - relates to pilot project rather than intervention itself. **Section 5 - Pregnancy section** **5.1. The visiting schedule** **5.1.1:** No relevant data **5.1.2 Clients' perceptions of the visiting schedule:** The study reports that clients of the programme generally agreed with the visiting schedule, regarding the weekly visits at the start of the programme as a way of getting to know the family nurse, and subsequent fortnightly visits as being 'just right, it's not too much time, just perfect' (p33). The study reports that all clients interviewed thought they had enough time with the family nurse. **5.1.3 Contact between visits:** The study reports that clients varied in the extent to which they contacted their family nurse between visits. Those who did contact the family nurse valued the 'between-session reassurance' (p34). 'If I ever worried I would always ask her and she's always there' (p34). **5.1.4 Nurses' experiences of programme delivery during the pregnancy period:** The study reports that family nurses find certain aspects of the pregnancy programme easier to deliver that others. The nurses reported that it was positive delivering the programme early on in clients' pregnancies. **Section 5.2 The content of contacts (p.37)** **5.2.1 Clients' perceptions:** The study reports that clients felt that the family nurse 'worked with them to decide what was talked about and regarded their nurse as flexible.' (p37). The study reports that clients felt they had enough time for each topic and 'if they wanted to talk more about something or still were not sure about a topic, they could ask to talk more at the next visit or, if necessary, request an additional visit' (p37). **5.2.2 Family nurses' perceptions:** The study reports that 'for the most part, family nurses found that they were able to use and manipulate the programme materials in ways that made each session relevant for their client' (p37). Nurses reported that the process of 'agenda-matching' enabled them to tailor the programme around each individual client. 'It's got to be what they're wanting, what's going on in their life at the time and let it flow from that' (family nurse, p38). **Section 5.3 The involvement of others**

	Section 5.3.1 Involvement of partners: The study found that, although clients were aware that they were able to bring a partner or parent to the visits, there were differing levels of involvement. Clients whose partners were not involved reported that it was because of work commitments, shyness on the part of the partner, or because the partner did not want to be involved. **Section 5.3.2 Family involvement:** The study found that involvement of family members in visits was generally limited. The study found that clients' mothers were very involved in some cases. This was reported to be positive by clients in that their 'mother's involvement in FNP had brought them closer together' (p39). **Section 5.4 Relationships between clients and family nurses:** The study reports that, in client interviews, Family nurses were described as 'really nice', 'a good laugh', 'funny' 'friendly' and 'great' (p40), and that clients' relationships with them were more like a friend than with a nurse or midwife. The study found that clients said they could be open and honest with their family nurse because they trusted them and knew that what they said would be confidential. Clients also reported that family nurses were easy to talk to, and perceived to 'have the time to sit and talk to clients in a way that midwives did not' (p.40). The authors conclude that 'clients' accounts thus provide evidence of the developing therapeutic alliance with their family nurse (p40). The study further reports that family nurses: - placed importance on building relationships and trust with clients, and that this was a foundation for being able to ask 'difficult questions' which might otherwise be perceived as intrusive - reported that they needed to strike a balance between giving information and 'sounding like a teacher' (p41). **Section 9. Discussion (p.64) Relevant conclusions to aspects of professional practice in early help:** 'The development of good, trusting relationships between family nurses and clients, with their descriptions of "agenda matching" in meetings and their views of their therapeutic relationships closely reflecting the central values and principles of the programme. The degree of trust and respect between clients and Family nurses also highlights the benefits of the strengths-based approach which underpins FNP in working with vulnerable young women' (p64). 'The influence of the family nurses' own degree of comfort in discussing sensitive issues on clients' willingness to discuss topics like sexual health' (p64).
Overall validity rating.	**Overall assessment of internal validity:** - Lack of information about sampling procedure and resulting participant sample is a serious flaw. **Overall assessment of external validity:** + Awarded + for external validity as only part of study is relevant to review question. **Overall validity rating:** -

	Little information is given with regard to how the participants in the qualitative research were sampled, what their characteristics are and how representative these are of the intervention population as a whole. Lack of information about participant characteristics (for both clients and family nurses) also makes it difficult to contextualise variation in research findings.

15. Paris R (2008) 'For the dream of being here, one sacrifices ...': Voices of immigrant mothers in a home visiting program. American Journal of Orthopsychiatry 78: 141–51

Research aims	**Study aim:** To capture the women's perception of the home visiting programme for mothers of infants and young children at risk of maltreatment services they were receiving. An additional aim of the study was to capture views on the process of immigration: these are not reported here. **Methodology:** Qualitative. **Country:** USA. **Source of funding:** Jessie Ball DuPont Fund and Community Benefits Programme, MA General Hospital.
PICO (population, intervention, comparison, outcomes).	**Participants:** Caregivers and families. Parents. **Sample characteristics:** • Age – 25–38 years of age, mean age 31 years. • Sex - mothers. • Ethnicity - Latino immigrant parents (11 from El Salvador, 1 from Honduras and 1 from Guatemala), been in the USA about 5 years. Spanish speaking. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - All 14 participants were unemployed. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - not relevant.

	• Intervention number - not relevant. • Sample size - 14 Latino mothers. **Intervention:** 'Visiting Moms' home visitation programme - a culturally sensitive programme using multilingual and bi-cultural para-professionals who were immigrants and mothers themselves. Typical services included weekly home visit and frequent phone contact, advocacy, parenting education and referral to resources. Services provided for up to 3 years. Weekly supervision by licensed social workers to mentor and teach home visitor's in the necessary relationship building skills and useful approaches to home visiting. **Outcomes measured:** Satisfaction with services - see narrative findings.
Findings.	**Narrative findings – qualitative and views and experiences:** Mothers' perception of home visitation service – **1. Emotional support:** Study found that respondents valued emotional support from home visitor, and participants reported that home visitor's listen, reassure, give guidance and link up appropriate agencies for assistance. 'When I really feel depressed … she would tell me to think things through ... she would counsel me and give me lots of advice' (p147). **2. Case management and advocacy:** Participants thought this was as important as emotional support, such as accompanying them to appointments, making phone calls, buying food/clothes, and referring to other services. 'She's helped me with food vouchers … filling out forms. I have called her and she has received me well … She's never denied me … toys for the kid … t's a big help' (p147). **3. Translation:** 13 of the 14 women in the study reported valuing the home visitor's role as translator and cultural brokers. **4. Teaching and friendship:** 6 of the women in the study described home visitor as 'educators'. Mothers learnt from home visitors about child development and the best ways to parent and the 'teaching' was informal. 'She tells me how to do things and if I cannot do it, she's there with me so that we can do it together'; 'She gives me advice about nutrition, food and a sleeping schedule for my child' (p148). Nine of the women interviewed saw home visitors as friends. **5. Mothers' dissatisfaction with the programme:** Mothers found transition to another home visitor difficult. It was also not appreciated when home visitors were inconsistent or unavailable. Data reported in the study on the following outcomes are not recorded here, as they are not directly relevant to research questions: Experiences of beginning of immigrant journey; life in the new country (USA). Trauma of immigra-

	tion, raising infants in a new country without familiar support and resources, isolation and managing challenging relationship with partners, etc. Mothers valued the emotional support from bilingual/bicultural home visitor and also the home visitor's role in advocacy and in providing practical help and guidance such as translation and dealing with various agencies. The home visitor's input in teaching parenting skills was also appreciated. Many mothers perceived their home visitor as a friend. Mothers found transition to another home visitor difficult, and were dissatisfied with the inconsistency and unavailability of some home visitors.
Overall validity rating.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** +

16. Self-Brown S, Frederick K, Binder S (2011) Examining the need for cultural adaptations to an evidence-based parent training program targeting the prevention of child maltreatment. Children and Youth Services Review 33: 1166–72

Research aims.	**Study aim: '**To assess the types of cultural adaptations that are being made to a widely implemented BPT, SafeCare, by providers working with families involved in the child welfare system, and to explore the need for more systematic adaptations ... to improve the program' (p1166). **Methodology:** Qualitative. **Country:** USA. **Source of funding:** Other - Doris Duke Charitable Foundation.
PICO (population, intervention, comparison, outcomes).	**Participants:** Professionals/practitioners - SafeCare providers who had experience of implementing SafeCare with diverse families and whose clients were always or often referred by child welfare agencies. **Sample characteristics:** • Age - Not reported. • Sex - Not reported. • Ethnicity - 5 Caucasian; 5 Latino; 1 African American. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported.

	• Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - not applicable. • Intervention number - not applicable. • Sample size - 11 SafeCare providers. **Intervention:** The context for this study is that of SafeCare, an evidence-based, BPT that targets risk factors for child physical abuse and neglect. SafeCare includes 3 modules: Health, Safety, and Parent–Child Interaction.
Findings.	**Narrative findings – qualitative and views and experiences:** The study reports the following findings – **1. Approaches to family engagement** Providers said that it is important to show openness and respect to gain trust during initial visits with the families: 'the main thing is just really get to know your clients and just let them be the expert and let them teach you … if you build a good relationship with your client that will be the key to how the implementation goes' (provider 11, p1168). Providers also reported that sometimes more than one visit was needed before starting training on SafeCare modules. '... spend one or more sessions in the engagement where you talk with [the families] about their concerns, engage the rest of the family … If you spend the time up front, it pays off drastically towards the rest of treatment' (provider 11, p1168). Matching providers and families by race/ethnicity was reported to be increasingly accepted, especially in matching based on language. Providers reported that families with a history with child protective services are often fearful of engaging in services, and suggested that this fear often increases for families who are of diverse backgrounds. Families who had been reported to child welfare services as a result of practices that were acceptable in their cultures found engagement with home visitors particularly difficult. ' ... some home remedies like sweating a fever out using cupping ... Things where in their cultures, it is not necessarily good … they do it as a last resort, but it's not a reportable offense where kids can be taken away' (provider 1, p1168). **2. The importance of flexibility in service delivery to retain families.** Providers reported that offering training sessions in different settings can improve family engagement: 'such as McDonalds, foster home/relative home, at the department … various locations as they try to find permanent housing' (provider 9, 10, p1168). Providers also noted that it is important to show flexibility about cultural or ethnic celebrations/traditions which can be disruptive to continuity of care: 'Even though a policy on consent around needing to maintain appointments, there are certain times where the parents are going to have Powwows or if there's a death, or things like that [resulting in many sessions being missed]

	… to be sensitive [about these events], which engages the parent a little bit more in the service, to the point that we can even double up on some sessions ...'(provider 10, p1168). **3. Perspectives on the SafeCare model** a. SafeCare delivered through home visitation Families appreciate home services; 'there are transportation issues, especially in rural areas; there's just not a lot there and so they love having somebody come to their home … I think they're very appreciative they don't have to worry about babysitting and everything else. They really view the workers as family friends and so I think it's been a very good approach' (provider 1, p1169). b. The focus on parenting skills providers reported that families enjoy SafeCare content, and particularly the focus on parenting skills. '… They just enjoy the feedback. Some of our parents who just feel a little insecure and really don't know if they're doing the right thing, it's important for them to have somebody that's there, that's observing … saying their doing well' (provider 7, p1169). c. Including children in sessions - Providers' perception was that inclusion of children in SafeCare was a positive. 'Parents like the idea of working with their children [during] a session … this gives them an opportunity to make the time and actually interact in a positive way with their children' (Latino family, provider 6, p1169). **4. Need for modifications or adaptations of structure or content of SafeCare sessions** Providers thought that a SafeCare programme adapted for particular culture/ethnic groups would be of limited use as every family is different: '[you cannot] just stereotype the family based on what [you] know about that culture, or assume that the family is the same as another family that was from the same culture' (provider 10, p1169). A more individualised approach to adaptation and case-by-case approach for specific local populations or specific families would be more appropriate. Learning about specific aspects of the cultures and beliefs of the populations was reported by providers to be important in establishing families' current beliefs about child rearing practices, medical treatment and so on. One example given was the importance of 'being open to discussing and working with home remedies, superstitions, and spiritual beliefs with families who consider these as a component of dealing with health issues is important … be respectful' (provider 1, p1169). Providers recommended adaptation of training materials/handouts to make them more comprehensive for families with language barriers and low literacy, such as adding picture aids and so on. **5. Participant recommendations** The study reports that participants made the following recommendations: - Adapt materials to be more user-friendly. - Provide additional training about cultural competency and cultural sensitivity to SafeCare home visitor and sharing information and experiences with other providers, especially related to delivering SafeCare to diverse families. The providers did not recommend systematic adaptations of the model for specific ethnic groups but provided general and specific information regarding SafeCare components that require adaptation on a case-by-case basis, which is likely to be applicable to many behavioural parenting programmes. Providers also emphasised the importance of flexibility in their approach; mutual respect, openness, avoiding stereotyping, in establishing a trusting relationship with the

	families, involving children and grandparents to improve a family's engagement with the programme. They also suggested making the training materials more user-friendly, tailored to meet the needs of the families from a diverse cultural background. Matching providers and families by race/ethnicity increasingly accepted, especially in matching based on language. Limitations small sample. This study is specific to the SafeCare Parenting programme, may not be applicable to all parenting programmes. No data from families. Uncertain if cultural adaptation would retain/maintain fidelity of the programme. Impacts of the cultural adaptation and changes made on the effectiveness of the intervention still unknown.
Overall validity rating.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** +

17. Stevens J, Ammerman RT, Putnam FW (2005) Facilitators and barriers to engagement in home visitation: A qualitative analysis of maternal, provider, and supervisor data. Journal of Aggression, Maltreatment and Trauma 11: 75–93

Research aims	**Study aim:** 'To explore reasons for and barriers to engagement in the first six months of parenting programs' (p76). **Methodology:** Qualitative – 5 focus groups: 2 mother groups, 2 home visitor groups, 1 supervisor group. **Country:** USA. **Source of funding:** Voluntary/charity - Every Child Succeeds (ECS), a voluntary home visitation programme.
PICO (population, intervention, comparison, outcomes).	**Participants:** • Caregivers and families. Mothers from a High Engagement Group (HEG) received at least 5 home visits and mothers from a Moderate Engagement Group (MEG) received no more than 2 home visits. • Professionals/practitioners. Home visitors with 6 months of home visitation experience, and, and home visitor supervisors; home visitation supervisors with at least 3 months experience of directly overseeing home visitors. **Sample characteristics:** • Age - HEG mothers: mean age 20.6 years old (SD=5.0), MEG mothers: mean age= 25.2 years (SD=5.4). • Sex - Mothers; all home visitors were female, all home visitor supervisors were female. • Ethnicity - HEG mothers: 5/8 African-Americans; MEG mothers: all (5) Caucasian HEG home visitors: 4/6 African American; MEG home visitors: 6/7 Caucasians. Home visitor supervisors: 6 Caucasian, 2 African-American. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported.

	• Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • Comparison numbers - Not applicable. • Intervention number - Not applicable. • Sample size - 8 mothers for HEG; 5 mothers for MEG. (total 13) Home visitors: 6 for HEG; 7 for MEG. (Total 13) home visitor supervisors: 8. **Intervention:** The context of this study is that of a voluntary home visitation programme in Midwest America called Every Child Succeeds (ECS).
Findings.	**1. Reasons for Participation:** The study reports key three reasons for participation 1a. Reassurance and non-judgmental support 1b. Information on effective parenting practices for mothers. '… learning how to discipline a child the correct way' (HEG mother, p83). 1c. tangible support (e.g., diapers) a recurrent theme among home visitors and mothers. Mothers anticipate these parent aid bags, toiletries, small toys, and children's books. Two additional themes related to regular maternal participation were: The flexibility and availability of home visitors. **2. Barriers to Participation:** The study reports two key themes in relation to barriers to participation: 2a. Invasiveness of programme due to inquiries about sensitive information (e.g., substance abuse, trauma history) during the initial face-to-face contacts. '… they need to redo those questions. You don't need to know all of my business. I just met you. Work on the relationship first, then you can ask me more in-depth questions' (MEG mother, p84). 2b. confusion amongst participating mothers regarding the frequency and benefits of home visits. 'I began to realize this is something that happens every week. Because I had no idea – I just thought people were coming to say hi!' (HEG mother, p84). '. . . And then they start not showing up or not being there. They never really understood what is was all about' (home visitor supervisor, p84). Two further barriers were reported by supervisors and home visitors: 2c. mental health problems which interfere with service delivery. ') 2d. challenges adapting the curriculum of the programme to fit the cognitive level of the individual mothers. The study reports that the following themes did not emerge consistently across groups, but did emerge in 1 or 2 groups: 2e. Concerns about being reported to child protective services was frequently mentioned in both mother groups. 2f. Home visitor's lack of knowledge, often attributed to a home visitor being childless herself. 2g. a perceived failure to receive helpful, accurate, and timely information appears to be a chief reason for disengagement. Mothers and home visitors perceived social support, psycho-education, and tangible assistance (receiving material

	goods such as diapers, toys and books for children) as key reasons for participation in home visitation programme. Mothers were concerned about the perceived invasiveness and intrusiveness of the assessments and the likelihood of them being reported to child protective services. Mothers who had relatively more resources would discontinue with the programme when the home visitor was perceived as offering too little useful and timely information. The mismatch between home visitor attributes and maternal needs was felt to be a barrier to continuous engagement with the programme. The availability and flexibility of the home visitor was also themes related to regular maternal participation. Many mothers would prefer visits by home visitors who were themselves parents. Home visitors and home visitor supervisors also faced challenging in working with mothers who had mental health problems; and in the need to adapt the programme to accommodate different cognitive needs of the mothers.
Overall validity rating.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** +

18. Voice Of Young People In Care (2014) Independent Inquiry on Child Sexual Exploitation in Northern Ireland: Consultation with care experienced young people. Belfast: Voice of Young People in Care

Research aims	**Study aim:** To inform an independent inquiry into child sexual exploitation in Northern Ireland by collecting the views of care experienced young people on this issue. 'Their understanding of child sexual exploitation (CSE). Where or from whom they learned about CSE. The ways in which a young person can be taken advantage of. The effectiveness of current safeguarding and protection arrangements. Measures being taken to prevent and respond to risks and CSE. Recommendations to prevent and respond to risk and CSE' (p9). **Methodology:** Qualitative study - workshops and questionnaires. **Country:** United Kingdom - Northern Ireland. **Source of funding:** Not reported. VOYPIC was asked to gather the views of care experienced young people in order to inform an independent inquiry which was supported by the Regulation and Quality Improvement Authority and the Criminal Justice Inspectorate of Northern Ireland.
PICO (population, intervention, comparison, outcomes)	**Participants:** Children and young people - Care experienced children and young people using services across 5 health and social care trusts in Northern Ireland. **Sample characteristics:** • Age - 12 to 25 years. 80% over 16 (n=44) 20% 12–15 (n=11).

	• Sex - Female 62% (n=36) Male 38% (n=19). • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - The study focuses on responses to risk of child sexual exploitation. • Looked after or adopted status - Non relative foster care n=8, Kinship care n=2, Children's home n=10, Independent living n=24, Secure accommodation n=9. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** N=55
Findings.	**Narrative findings – qualitative and views and experiences:** The authors note that young people perceived there to be different responses to incidents and risks depending on where the child was placed. Participants were reported to believe that '… current measures to prevent and respond to risks and CSE are not working as effectively as they could and that a better approach should be considered' (authors, p22). The young people discussed responses to risk and the authors identify three themes which emerged – differing responses used by foster carers and children's home staff; a perception that responses were often an 'overreaction' and that police were sometimes unnecessarily involved; and views on the use of secure accommodation. However, as the decision to remove children and young people from their current home is out of scope, findings in relation to this have not been extracted. The authors report that some young people felt that foster carers responses '… to risky behaviour is effective' (authors, p23). They note that foster carers used boundaries to resolve problems and that calm and continuous discussion were an important feature. This was contrasted with responses from staff in children's homes, where young people noted that staff turnover made it more difficult to build relationships which in turn made it difficult for staff and young people to discuss their behaviour and the appropriate response. They also report that young people felt that the lack of confidentiality in children's homes was a barrier to approaching staff. The authors note that '… nearly every group …' (authors, p24) commented on perceived overreactions by staff in children's homes and the tendency for '… excessive contact …' with the police when a young person was not at home when expected. The authors also report that young people felt that applying 'sanctions' and curfews to all children living in a particular home was unfair.

	The authors report that participants were aware of 'Harbouring Orders' but had little understanding of how these actually worked. The authors conclude that '… excessive contact …' with the police is ineffective and '… may result in a young person refusing to engage with support staff or the police service ...' (authors, p27). They also state that excessive questioning and the use of 'sanctions' such as less pocket money are of '… limited value or effect. Instead, young people felt that staff should talk to them more and highlight the consequences of the risks they may be taking by using drugs and alcohol and agree together acceptable behaviour and how to keep safe' (authors, p28). Participants felt that children's home staff relied more heavily on the police due to a lack of confidence: 'Staff don't have confidence which is why police is phoned [sic]' (participant, p30). The authors report that some young people felt that this could be addressed through training and suggest that training on responding to risk would be most effective if it involved children and young people. Young people were reported to believe that the organisation of activities in children's homes and encouragement to volunteer was a potential means of minimising risky behaviour, with one young person stating that: 'Getting young people involved in the local community and in volunteering … with VOYPIC, youth club … if you have too much time on your hands, you can end up hanging out with the wrong crowd' (p33). Young people also felt that if staff spent more time with them the same effect could be achieved. The authors report that young people felt that peer education and support and the use of real life examples would help to educate them on their vulnerability in relation to child sexual exploitation. In their conclusions the authors note that participants did not discuss safety plans.
Overall validity rating.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall validity rating:** - The study is very unclear on a number of key methodological issues such as sampling techniques, analysis procedures and the contexts in which data were collected.

19. Woodman J, Gilbert R, Allister J et al. (2013) Responses to concerns about child maltreatment: A qualitative study of GPs in England. BMJ Open 3: e003894

Research aims	**Study aim:** To explore '... how a small sample of GPs understood and responded to child maltreatment-related concerns in their daily practice' (p2).

	Methodology: Qualitative study - In-depth, individual interviews conducted in person. **Country:** UK - England. **Source of funding:** Other - MRC/ESRC interdisciplinary studentship award (grant number G0800112).
PICO (population, intervention, comparison, outcomes).	**Participants:** • Professionals/practitioners - GPs, practice nurses and health visitors from English GP practices. The number of years which the GPs had been practicing for ranged between 5 and 40 (average of 19). They had been at their current practice between 6 months and 23 years (average of 10). **Sample characteristics:** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - The study aims to understand GP responses to children for whom there are maltreatment concerns but do not meet social care thresholds for intervention. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** • GPs – n=14. • Practice nurses - n=2. • Health visitors – n=2.
Findings.	**Narrative findings – qualitative and views and experiences:** **A - Which families do GPs intervene with?** The authors identify 3 themes which were '... typical of accounts of intense or long-term involvement with maltreatment-related concerns' (p4). These were: 1. GPs tended to conceptualise their responses as 'medical' and were able to justify their safeguarding involvement

	with families who had higher medical needs on this basis. They suggest that this '... containment of safeguarding within a medical sphere seemed most compatible, with chaotic, neglectful families seen to be suffering a host of medical and social problems' (authors, p4). The authors report that a number of GPs felt that intervening in cases where there were maltreatment concerns without any apparent medical need was not part of the GPs responsibility. However, other GPs challenged the distinction between 'health' and 'social' need. 2. GPs appeared to be more willing to intervene in cases where the parents were viewed as 'incompetent' rather than 'malicious' (p4). 3. GPs also appeared to be more willing to take action when they expressed distrust regarding the involvement of social services who were perceived to respond inappropriately. **B - Actions taken:** The authors identify seven types of actions which GPs described taking in relation to maltreatment-related concerns: 1. Monitoring concerns (e.g., 'using routine health checks in children and regulator consultations for health problems in parents to assess well-being of children and coping/risk factors in parents). 2. Advocating (e.g., support requests for improved housing). 3. Coaching parents (e.g., talking to parents, usually the mother, to encourage them to change their behaviours) 4. Providing opportune healthcare for children (e.g., using a consultation for something else to deliver an overdue vaccination). 5. Referral to other services (e.g., referral to Children's Social Care). 6. Working with other services. 7. Recording the concerns. The authors note that GPs '... were very aware their management of maltreatment-related concerns relied on regular contact with families for nonmaltreatment related reasons (monitoring and opportune healthcare), help-seeking behaviour and honest disclosure of problems from adult family members (monitoring and advocating), parental engagement with general practice (coaching and advocating) and being able to offer services that parents wanted (monitoring and opportune healthcare)' (authors, p5). **C - Facilitators and barriers:** 1. Relationship between GPs and families - The authors report that GPs attempts to foster trust between themselves and parents was '... the strongest and most persistent theme across the interviews ...' which was seen as an important facilitator to response and a means of encouraging parents to engage and accept help and advice (p5). One GP reported that: 'It's [the reason to develop trust] not frightening them away because, as well, there is that kind of unseen agreement between you. She is thinking: 'if this gets a bit much for me, I might be asking you for a bit more help'. 'How will you be when I ask you for more help?' and I am thinking 'if this gets too much for you I might ask you if you need more help. I want you to be accepting of that help and not worried about it' (participant 0, discussing a 4 year old child

	with older siblings). GPs felt that it was easier to foster trust when they could offer something to families such as a letter to support a benefit or housing claim. 2. Relationship between GPs and health visitors (p5) - The study found that GPs reported being dependent on health visitors in their responses to maltreatment-related concerns. The authors report that the two health visitors interviewed believed that GPs had only limited knowledge and were eager to avoid or pass on child protection work to other professionals. 'I think ultimately being based in the same building, seeing people day to day, you know in the kitchen, putting the kettle on, that kind of daft thing does build a good relationship' (participant 16, discussing siblings aged 3 and 7 years old) (p 8). In their discussion the authors also report that '... information and support from health visitors was threatened by mismatched expectations ...' (p9). 3. Relationship between GPs and other professionals - GPs did not discuss in detail how their relationships with other professionals helped or hindered their responses to maltreatment concerns. However, they '... wished to be seen as separate from children's social care and paediatric services, which they thought patients saw as punitive and policing' (p5). The authors describe a lack of feedback from children's social services, seen to be exacerbated by lack of personal relationships between GPs and social workers, and note that GPs drew on personal contacts with 'trusted' paediatricians. 4. Medical role GPs justified and legitimised their involvement by framing their responses as 'medical'. The authors note that 'the theoretical distinction between 'medical' and 'social' problems was used by participants to delineate where the GP could legitimately be involved with maltreat-related concerns. However this neat distinction was challenged' (p8). **Discussion:** The study found that GPs '… described being actively involved with the management of (possible) child neglect and emotional abuse' (p9). Study notes that, due to 'case-based' study design, it is possible that some GPs recounted what they should have done, rather than what they did do. It is also unclear to what extent the 7 responses are being used in general practice more widely. Study findings are compared to a study by Tompsett et al. (2010). The authors note that, similar to this study, Tompsett et al.'s study suggests that GPs might have the biggest role to play for children with chronic neglect. The authors note that this study did to seek the views of parents or children, which is a limitation to the study. Implications The authors suggest that the implications of the study include: - Policy and research focus should be broadened to include direct intervention by GPs for families who prompt maltreatment-related concerns.
Overall validity rating.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++

	Overall validity rating: +

Response

Review question 15 – What is the impact of social and psychological interventions responding to child abuse and neglect? (Prevention of recurrence, prevention of impairment)

Review question 15 – Critical appraisal tables

1. Barlow J, Johnston I, Kendrick D et al. (2006) Individual and group-based parenting programmes for the treatment of physical child abuse and neglect. Cochrane Database of Systematic Reviews issue 3: CD005463

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. **Adequate description of methodology?** Yes. **Rigorous literature search?** Yes.	**Inclusion of relevant individual studies?** Yes. **Study quality assessed and reported?** Yes. **Do conclusions match findings?** Yes.	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Not reported. **Were service users involved in the study?** Not reported. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes - parents and families.	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** + **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Is the study setting the same as at least 1 of the settings covered by the guideline? Yes - clinic and home-based. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes – response, treatment of physical child abuse and neglect. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **Does the study have a UK perspective?** No – mainly US studies.	

2. Browne DT, Puente-Duran S, Shlonsky A et al (2016) A Randomized Trial of Wraparound Facilitation Versus Usual Child Protection Services. Research on Social Work Practice 26: 168-179

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim Study aim: To evaluate whether the addition of a wraparound facilitator to regular child protection services improved child and family functioning over 20 months.	**Was the exposure to the intervention and comparison as intended?** Partly. It is likely that workers in the control condition (usual CPS care) also ‘‘came up with new ideas when others weren’t working’ (similar to that of wraparound care) (p7). **Was contamination acceptably low?**	**Does the study’s research question match the review question?** Yes. To evaluate whether the addition of a wraparound facilitator to regular child protection services improved child and family functioning over 20 months. **Has the study dealt appropriately with any ethical concerns?**	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** + **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Description of theoretical approach? Yes **How was selection bias minimised?** Randomised: Blocked randomization stratified by site (3 sites), with variable block sizes of 2, 4, and 6. The unit of analysis and randomization was at the family level, but only one child per family was included. **Was the allocation method followed?** Yes: The allocation ratio was 1 control: 1 intervention. Random numbers generated and placed in opaque, sealed envelopes. **Is blinding an issue in this study?** Part blinding. Single blinding: interviewers were blind to the family's experimental condition.	Partly. Commonalities (community-based care and unconditional care) and differences in experimental and control conditions likely as is often the case in psychosocial intervention (pp7, 10). **Did either group receive additional interventions or have services provided in a different manner?** Not reported. **Were outcomes relevant?** Yes. Child impairments, caregiver psychological distress. **Were outcome measures reliable?** Yes. Validated. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. At 20 months after intervention. **Was follow-up time meaningful?** Yes.	Yes. Ethical approval obtained from the McMaster University Ethics Review Board, plus procedures for informed consent. **Were service users involved in the study?** Yes. Service users as participants in the study. **Is there a clear focus on the guideline topic?** Yes. Impact of social intervention on child protection. **Is the study population the same as at least one of the groups covered by the guideline?** Yes. Family care givers (mostly mothers). **Is the study setting the same as at least one of the settings covered by the guideline?** Yes. Not reported, but likely to be in home setting. **Does the study relate to at least one of the activities covered by the guideline?** Yes. Child protection (abuse and neglect).	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Did participants reflect target group? Yes. Families who had a substantiated investigation for child maltreatment. **Were all participants accounted for at study conclusion?** Partly. Initially 306 families were approached and 135 agreed to participate in the study prior to being randomised into intervention and control group. Demographics and reason for referral/maltreatment type between the participants (n=135) and non-participants (n=171) were similar, except that significantly more of the 25–29 aged group declined to take part (non -participants) (10.4% vs. 24%, p=0.02). Significantly more participants had financial problems (30.4% vs. 18.1%, p=0.02) and previous child removal due to CPS involvement (26.7% vs.18.1%, p=0.01). 16 of the initial 135 families	**Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Partly. Both groups similar at baseline except for the following: Parents in intervention group more likely than control to have been reported for emotional harm/exposure to conflict (41% vs. 22%, p=0.02) and have substance use problems (21.7% vs. 3.4%, p=0.003). Fewer parents in the intervention group had spouses who were employed (16.7% vs. 39%, p=0.03) and had substantial unmet material needs (10% vs. 50%, p=0.05) (p7). **Was intention to treat (ITT) analysis conducted?** Yes. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Yes. Power calculation performed, power analyses on all outcomes were deemed to be satisfactory (>0.80) (p6) **Were the estimates of effect size given or calculable?** Yes.	**(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **Does the study have a UK perspective?** No. Canada.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
were lost to follow-up (attrition rate 11.9%). Completers and noncompleters were similar on all other variables at baseline, except that slightly greater proportion of children among completers were removed from the home due to previous CPS involvement (27% vs. 25%, p=0 .02); a lower proportion of retained caregivers were involved in adult conflict/violence (16% vs. 43.8%, p=0 .04); Retained children were lower in baseline behavioural and emotional strengths scores (p=0 .03), and in the fire setting subscale on the CAFAS (4% vs. 28.6%, p=0.01).	**Were the analytical methods appropriate?** Yes. Intention-to-treat principle, analysis by ANOVA using SPSS. To adjust for attrition, data were imputed using multiple imputation. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. **Do conclusions match findings?** Yes.		

3. DePrince AP, Chu AT, Labus J S et al (2015) Testing Two Approaches to Revictimization Prevention Among Adolescent Girls in the Child Welfare System. Journal of Adolescent Health 56: S33-S39

Internal validity - approach and sample.	Internal validity - performance and analysis.	External validity.	Overall validity rating.
Study aim Study aim: To 'compare two interventions designed to decrease revictimisation	**Was the exposure to the intervention and comparison as intended?**	**Does the study's research question match the review question?** Yes.	**Overall assessment of internal validity:** +

Internal validity - approach and sample.	Internal validity - performance and analysis.	External validity.	Overall validity rating.
in a diverse sample of adolescent child-welfare involved girls' (pS33). **Description of theoretical approach?** Yes. Testing two different theoretical approaches: social learning/feminist theory compared to risk detection and executive function. **How was selection bias minimised?** Randomised. Although non-treatment group was not randomly selected. Results from this group are not presented. **Was the allocation method followed?** Yes. **Is blinding an issue in this study?** No blinding. Participants not blind to study condition. Unclear whether researchers administering outcome measures were	Partly. Attendance at sessions was 73% for risk detection/executive function intervention and 70% for social learning/feminist theory. **Was contamination acceptably low?** Yes. **Did either group receive additional interventions or have services provided in a different manner?** No. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Yes. Standardised scales used - Traumatic events screening inventory (TESI) and Conflict in Adolescent Dating Relationships (CADRI). Reliability of scales not reported. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Partly. Measurement did not look at impact on the girls' overall wellbeing.	**Has the study dealt appropriately with any ethical concerns?** Partly. Ethical approval obtained. Informed consent from parents/guardians, although unclear if consent obtained from young people themselves. **Were service users involved in the study?** No. Service users involved as participants only. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least one of the groups covered by the guideline?** Yes. Females aged 12 to 19 who had histories of childhood neglect or abuse. **Is the study setting the same as at least one of the settings covered by the guideline?** Yes. **Does the study relate to at least one of the activities covered by the guideline?** Yes. Study relates to response.	**Overall assessment of external validity:** ++ US study, but background services likely to be the same for this type of intervention as in UK. **Overall validity rating:** + Limitations include no intent to treat analysis, and creation of a 'no treatment' comparison group of those who did not attend any sessions.

Internal validity - approach and sample.	Internal validity - performance and analysis.	External validity.	Overall validity rating.
blind to participant condition. **Did participants reflect target group?** Yes. **Were all participants accounted for at study conclusion?** Yes.	**Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Partly 4 months. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. Authors conclude two intervention groups were similar at baseline. **Was intention to treat (ITT) analysis conducted?** No. No intent-to-treat analysis conducted. Some young people who were initially randomised to a treatment group, but attended no sessions, were moved in to a 'no treatment' comparison group. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Partly. No power calculation given, but relatively large sample size. **Were the estimates of effect size given or calculable?** No.	**(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **Does the study have a UK perspective?** No. US study.	

Internal validity - approach and sample.	Internal validity - performance and analysis.	External validity.	Overall validity rating.
	Were the analytical methods appropriate? Yes. **Was the precision of intervention effects given or calculable? Were they meaningful?** No. **Do conclusions match findings?** Yes.		

4. Donohue B, Azrin NH, Bradshaw K et al. (2014) A controlled evaluation of family behavior therapy in concurrent child neglect and drug abuse. Journal of Consulting and Clinical Psychology 82: 706–20

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim 'The study examines 'the effects of Family Behavioral Therapy as compared to treatment as usual community-based services (TAU)' (p708). **Description of theoretical approach?** No. No logic model provided for how intervention operates. **How was selection bias minimised?** Randomised. –Urn randomisation.	**Was the exposure to the intervention and comparison as intended?** Yes – Fidelity monitored during intervention. 95% of protocol instructions were implemented by providers. **Was contamination acceptably low?** Yes. –No contamination between groups reported. **Did either group receive additional interventions or have services provided in a different manner?**	**Does the study's research question match the review question?** Yes – Study evaluates the impact of an intervention for the families of children who are experiencing or have experienced abuse and neglect. **Has the study dealt appropriately with any ethical concerns?** Yes. –Approval obtained from institutional review board and informed consent obtained from all participants.	**Overall assessment of internal validity:** **-** **Overall assessment of external validity:** **++** **Overall validity rating:** **-** A key methodological flaw is the failure to report cell sizes for drug-exposed versus non-drug exposed families. This makes it difficult to judge the validity of the statistical analysis. There is also lack of clarity regarding data imputation methods for intent to treat analysis (or indeed if imputation was used).

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Was the allocation method followed? Yes. **Is blinding an issue in this study?** Part blinding. –Not possible to blind participants or providers to study condition. However, assessments of outcome were carried out by providers who were blind to study condition. **Did participants reflect target group?** Partly –Requirement of intervention is that there must be one other adult prepared to participate in the parent's treatment (p708). This could potentially skew the data in favour of those with better support networks. The authors do not comment on the representativeness of the sample compared to target group. **Were all participants accounted for at study conclusion?** Yes –All participants accounted for (See Fig 1, p710) but significant attrition, with attrition from assignment to second follow up 26% in	No. –Although TAU option could comprise a variety of services. **Were outcomes relevant?** Yes. –Focusing on parenting and parent wellbeing outcomes relevant to substance misuse. **Were outcome measures reliable?** Yes. **Were all outcome measurements complete?** Yes. –Primary outcome measures (risk of abuse, frequency of days using hard drugs or marijuana) taken at baseline and 6- and 10-month follow up. Secondary outcome measure (risk of HIV transmission) taken at 6- and 10-month follow-up only. **Were all important outcomes assessed?** Partly. –No outcome measures relating to child wellbeing were included. **Were there similar follow-up times in exposure and comparison groups?**	**Were service users involved in the study?** No. –Service users involved as participants only - not in design or interpretation of findings. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. –Families who have been referred to child protective services on grounds of child neglect and treatment for substance misuse. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. –Intervention delivered in the home. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. –Response to prevent recurrence and prevent/ameliorate impairment.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
experimental group, and 14% in control group.	Yes. **Was follow-up time meaningful?** Partly –Follow-up time sufficient to observe proximal impact on outcomes, but not longer term outcomes. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. –Analysis of demographic and outcome variable at baseline using chi square and ANOVA showed no significant baseline differences between participants in the intervention versus the comparison group. **Was intention to treat (ITT) analysis conducted?** Partly. –The study states that 'all 72 of the qualifying participants who were interested in participating in the study were randomly assigned to treatment (35 FBT, 37 TAU) and included in the intent to treat study analyses' (p709). However, as only 55 people provided data at 6 months, and 58 provided at 10 months and no method for imputing missing data is reported.	**(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **Does the study have a UK perspective?** No. –US study, however the service delivery system (CPS making referral to a specialist provider) is similar to UK.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Was the study sufficiently powered to detect an intervention effect (if one exists)? Partly. –Power calculation not given. Sample size is relatively small, particularly given the 2x2 (x3) design, meaning that sample size in each condition would be relatively small (c.18). **Were the estimates of effect size given or calculable?** Yes. –Effect sizes calculated using partial eta squared. **Were the analytical methods appropriate?** Yes. –Analysis of variance with no covariates (no existing differences identified between intervention and control group). **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. **Do conclusions match findings?** Yes.		

5. Fantuzzo J, Manz P, Atkins M et al. (2005) Peer-mediated treatment of socially withdrawn maltreated preschool children: Cultivating natural community resources. Journal of Clinical Child and Adolescent Psychology 34:320–5

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim –To evaluate the effectiveness of Resilient Peer Treatment, a '… peer-mediated, classroom-based intervention for socially withdrawn, maltreated pre-school children' (p320). The intervention had previously been found to be effective in a treatment setting and the authors aimed to determine whether this could be transferred to the classroom setting. **Description of theoretical approach?** No. –The authors do not present a clear theory of change or logic model but note that the intervention is designed to improve social functioning by providing an opportunity for '... positive play experiences with peers, who evidence high social functioning amidst high-risk urban contexts' (p321). **How was selection bias minimised?** Randomised. –Method not reported.	**Was the exposure to the intervention and comparison as intended?** Yes. –The authors report that checklists were used to determine whether the interventions had been carried out as planned. This is reported as an average of 90%, ranging from 81–100%. They also note that sessions were not conducted when either the 'play buddy' or participating child was absent. **Was contamination acceptably low?** Not reported. –Information on contamination is not provided. **Did either group receive additional interventions or have services provided in a different manner?** No. –There is no indication that either group received additional interventions. **Were outcomes relevant?** Yes.	**Does the study's research question match the review question?** Yes –The study aimed to evaluate the effectiveness of Resilient Peer Treatment, a '… peer-mediated, classroom-based intervention for socially withdrawn, maltreated pre-school children' (p320). The intervention had previously been found to be effective in a treatment setting and the authors aimed to determine whether this could be transferred to the classroom setting. **Has the study dealt appropriately with any ethical concerns?** Partly. –Informed consent and approval for the study are not reported but the authors note that 'permission' was sought prior to randomisation. **Were service users involved in the study?** No. –In discussing the context for the intervention, the authors note that they '… initiated a partnership pro-	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall validity rating:** - The use of coding systems and scales with unclear reliability and validity and a very short follow-up (two weeks) mean that it is difficult to be confident in the findings of this study.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Was the allocation method followed? Not reported. –Method of allocation and concealment are not reported. **Is blinding an issue in this study?** Part blinding. –Due to the nature of the intervention it would not have been possible to blind participating children, 'Play buddies' or 'play supporters' however the authors note that observational data was coded by researchers blind to maltreatment status, group assignment and assessment point. They also report that teachers who provided data on children's play and behaviour in the classroom were blind to maltreatment status and group assignment. **Did participants reflect target group?** Partly. –The number of eligible children and those for whom consent for participation was provided is not reported. Eligibility was determined on the basis of teacher ratings of prosocial peer interactions and independent verification of	–The study aimed to evaluate the effectiveness of an intervention for socially withdrawn maltreated preschool children and the outcome measures used related to social behaviours and interactions with peers. **Were outcome measures reliable?** Partly. –Only two of the scales or coding systems appear to have established reliability and validity and data in relation to this is only provided for 1 measure. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** No –Follow-up assessments were conducted at the two week post-	cess with Head Start staff and parents …' (p321) but the extent to which parents were involved in the design of the intervention or the study is not clear. **Is there a clear focus on the guideline topic?** Yes. –The study aims to determine the effectiveness of an intervention designed for maltreated children. **Is the study population the same as at least 1 of the groups covered by the guideline?** Partly. –Participants were socially withdrawn African American children enrolled in the Head Start programme. Child maltreatment was substantiated for n=37 out of n=82. The study has been included in the NCCSC review because findings are reported on the basis of treatment group and maltreatment status. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes –The setting for the intervention and assessments was the child's classroom.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
these. The entire sample was African American. **Were all participants accounted for at study conclusion?** Yes. –There were no participants lost to follow-up or and none failed to complete the programme.	intervention point which is extremely short and unlikely to allow any effects of the intervention to be detected. **Analyses** **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Not reported. –Pre-treatment comparisons focused on comparing maltreated versus non-maltreated children, rather than assessing any differences between exposure and comparison. **Was intention to treat (ITT) analysis conducted?** Yes. –All participants completed the intervention and participated in follow-up assessments. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported.	**Does the study relate to at least 1 of the activities covered by the guideline?** Yes. –Response – The study evaluates an intervention which is designed to enhance the social capabilities of maltreated pre-school children who are withdrawn. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. –Outcomes included social interactions with peers and social behaviour. **Does the study have a UK perspective?** No. –Study was conducted in the USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	–The authors do not provide a power calculation or expected effect sizes. The sample size is acceptable. **Were the estimates of effect size given or calculable?** Yes. –Partial eta squared effect sizes are reported. **Were the analytical methods appropriate?** Yes –Analysis of variance and chi-square analysis. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. –p values and partial eta squared effect sizes are provided. **Do conclusions match findings?** Partly. –On the whole the authors conclusions fit well with their findings, however, it should be noted that some of the findings for which a significant effect was found were observed within a setting which overlaps with the intervention under evaluation (i.e. observations of		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	children's play in dyadic play corner interactions with their play buddy).		

6. Fisher PA, Nurraston B, Pears KC (2005) The Early Intervention Foster Care Program: permanent placement outcomes from a randomized trial. Child Maltreatment 10:61–71

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim This paper forms part of a larger RCT exploring effectiveness of Early Intervention Foster Care Program. This paper reports permanent placement outcomes. **Description of theoretical approach?** Partly. Some background to the intervention given, but paper also refers to Fisher et al. 1999, 2000). **How was selection bias minimised?** Randomised. **Was the allocation method followed?** Yes. **Is blinding an issue in this study?** No blinding.	**Was the exposure to the intervention and comparison as intended?** Yes. **Was contamination acceptably low?** Yes. **Did either group receive additional interventions or have services provided in a different manner?** Not reported. **Were outcomes relevant?** Partly. Study also explored a number of cognitive and behavioural measures, but they are not reported in this paper.	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Partly. No reference made to ethical approval. Consent sought from caseworker (legal guardian) and, where possible, birth parents. No mention of obtaining informed consent from children. **Were service users involved in the study?** No. Service users involved as participants only. **Is there a clear focus on the guideline topic?** Yes.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Unclear to what extent US training for foster carers is similar to UK. **Overall validity rating:** + Relatively small sample size (n=54). Focuses on placement outcomes only.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Participants not blind to intervention condition. Not reported whether those assessing outcomes were blind to intervention condition. **Did participants reflect target group?** Yes. **Were all participants accounted for at study conclusion?** Yes.	**Were outcome measures reliable?** Yes. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Partly. As above - this paper focuses on placement outcomes. Other types of outcomes were also measured but are not reported here. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Yes. 24 months. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Partly. Raw data provided suggest that there are some key differences in terms of type of permanent placement (EIFC is 48% birth parents, RFC 68%) which may have affected results.	**Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Partly. Placement outcomes will be coded as a service outcome. **Does the study have a UK perspective?** No. US study.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Was intention to treat (ITT) analysis conducted? No. Eight children did not complete the assessments, and missing data do not appear to have been imputed. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. No power calculation given, although effective sample size is relatively low (n=54). **Were the estimates of effect size given or calculable?** No. **Were the analytical methods appropriate?** Yes. **Was the precision of intervention effects given or calculable? Were they meaningful?** No. **Do conclusions match findings?** Yes.		

7. Forrester D, Holland S, Williams A et al. (2014) Helping families where parents misuse drugs or alcohol? A mixed methods comparative evaluation of an intensive family preservation service. Child & Family Social Work 21(1) 65-75

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
Qualitative comp 1 Which component? Family's experience of using Option 2, the experience of other services; and family life and experiences since the time of referral (data collected via interviews). Most participants also gave extraordinary stories, 'autobiographical narratives' placing and explaining their difficulties that ran from before substance misuse became a problem, up to the time of the interview (p69). **1.1 Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. Family life and experiences since the time of referral, also 'autobiographical narratives'.	**3. Quantitative component (incl. non-RCT; cohort study; case-control study)** Which quantitative component? Quasi-experimental (Comparative cohort) study: Option 2 vs. no Option 2 (control). **3.1 Are participants (organisations) recruited in a way that minimises selection bias?** No. Initially a retrospective study, intended to match each comparison family to the O2 family referred for the same level of seriousness but due to a low response rate (36%) when 75 families were approached, the case comparison approach was dropped and more recent O2 referrals were approached to take part in the research. This compromises the comparability of the samples (p67). **3.2 Are measurements appropriate (clear origin, or validity known, or standard instrument; and absence of contamination between groups when appropriate) regarding the exposure/intervention and outcomes?** Yes. Instruments used: 1. Child's welfare (Strengths and Difficulties Questionnaire [SDQ] for emotional and behavioural development; Goodman 2001) 2. Parental	**Does the study's research question match the review question?** Yes. Child protection relating to parental substance abuse. **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval given by the University of Bedfordshire Research Ethics Committee. All details anonymized. **Were service users involved in the study?** Yes. Service users as participants in study. **Is there a clear focus on the guideline topic?** Yes. response to child protection concerns **Is the study population the same as at least one of the groups covered by the guideline?** Yes. Parents and families. **Is the study setting the same as at least one of the settings covered by the guideline?**	**Overall assessment of internal validity** - **Overall assessment of external validity** ++ **Overall validity score** - Unclear if the groups in the quasi-experimental cohort study are comparable.

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
1.2 Is the process for analysing qualitative data relevant to address the research question? Yes. To explore parents' experience of the Option 2 programme. From the stories and the 'autobiographical narratives' describing the parents' life journeys, the impact of the O2 program was assessed. **1.3 Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** Unclear. Not reported. **1.4 Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?** Unclear. Not reported	substance use (Maudsley Addiction Profile [Section B]; Marsden et al. 1998) 3. Parental emotional well-being (General Health Questionnaire [GHQ-12]; Goldberg 1978) 4. Family functioning (Family Environment Scale [FES subscales for family cohesion, open expression of emotion and open conflict]; Moos & Moos 1986). **3.3 In the groups being compared (exposed versus non-exposed; with intervention versus without; cases versus controls), are the participants comparable, or do researchers take into account (control for) the difference between these groups?** Partly. At time of referral, no significant difference between the O2 group and control in 1. Age of the parents. 2. Age of the children. 3. Number of children living with their parents. 4. Number of families who accepted referral to avoid care. There were 1. Significantly more parents with illicit drug use in the O2 group (72% vs. 23%, p=0.01) (Note Table 1 shows =0.1. However, as all other p values > 0.05 are shown as NS, and due to large difference in proportions, assume this is a typo and should read p=0.1). 2. Significantly higher proportion of boys in the Option 2 group (70& vs 43%, p=0.02).	Yes. Not reported, but assumed to be home setting. **Does the study relate to at least one of the activities covered by the guideline?** Yes. Child protection. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes.	

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
	3.4 Are there complete outcome data (80% or above), and, when applicable, an acceptable response rate (60% or above), or an acceptable follow-up rate for cohort studies (depending on the duration of follow-up)? Unclear. **5.1. Is the mixed-methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and quantitative aspects of the mixed-methods question?** Yes. **5.2. Is the integration of qualitative and quantitative data (or results) relevant to address the research question?** Yes. **5.3 Is appropriate consideration given to the limitations associated with this integration, such as the divergence of qualitative and quantitative data (or results)?** Yes There were several steps taken in the analytical process of these 'stories'. All transcripts and summaries were discussed among the researchers to ensure agreement, then entered into NVivo, coded and		

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
	analysed. The authors also highlighted the limitations of combining the qualitative and quantitative data in a study which was compromised by small sample size and group comparability due to failure to recruit sufficient no. of participants (p72).		

8. Goldman Fraser J, Lloyd SW, Murphy RA et al. (2013) Child exposure to trauma: Comparative effectiveness of interventions addressing maltreatment. Comparative Effectiveness Review 89: 1–161

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. **Adequate description of methodology?** Yes. **Rigorous literature search?** Yes.	**Inclusion of relevant individual studies?** Yes. **Study quality assessed and reported?** Yes. **Do conclusions match findings?** Yes.	**Does the study's research question match the review question?** Partly. –Include sexually abused children. **Has the study dealt appropriately with any ethical concerns?** Not reported. **Were service users involved in the study?** Not reported. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes.	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** + **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Is the study setting the same as at least 1 of the settings covered by the guideline? Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. –No data on abuse recurrence. **Does the study have a UK perspective?** No. US, UK, Canada.	

9. Graham-Bermann SA, Miller-Graff LE, Howell KH et al (2015) An Efficacy Trial of an Intervention Program for Children Exposed to Intimate Partner Violence. Child Psychiatry and Human Development 46: 928–39

Internal validity - approach and sample.	Internal validity - performance and analysis.	External validity.	Overall validity rating.
Study aim Study aim: 'To compare outcomes for 4-6 year old children randomly assigned to a program designed to address the effects of exposure to IPV with those allocated to a waitlist comparison condition' (p928). **Description of theoretical approach?** No. **How was selection bias minimised?** Randomised. Quasi-RCT, as participants were alternately allocated in blocks of 5- first 5 to intervention and next 5 to control (wait list). **Was the allocation method followed?** Partly. Mothers were assigned to the experimental condition by the project coordinator. The first five families who qualified for the study were assigned to	**Was the exposure to the intervention and comparison as intended?** Not reported. **Was contamination acceptably low?** Not reported. **Did either group receive additional interventions or have services provided in a different manner?** Not reported. **Were outcomes relevant?** Yes. Child internalizing disorders. **Were outcome measures reliable?** Yes. The Child Behavior Checklist. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?**	**Does the study's research question match the review question?** Yes. Children exposed to IPV (intimate partner violence). **Has the study dealt appropriately with any ethical concerns?** Yes. University Institutional Review Board approval obtained, interviewers trained in research ethics. Mothers gave informed consent prior to being interviewed, and were compensated $25 for each study interview. **Were service users involved in the study?** Yes. Mothers and child pairs were participants in the trial. **Is there a clear focus on the guideline topic?** Yes. IPV (intimate partner violence). **Is the study population the same as at least one of the groups covered by the guideline?** Yes. Children exposed to IPV (intimate partner violence). **Is the study setting the same as at least one of the settings covered by the guideline?**	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** +

the experimental condition, and the next five families were assigned to the no treatment comparison condition. All participating agencies used the same allocation schedule. **Is blinding an issue in this study?** Blinding. Participants and those who conducted interviews were blind to group assignment. **Did participants reflect target group?** Yes. Mother-Child pairs who experienced IPV. **Were all participants accounted for at study conclusion?** Yes. Study completion rate (drop-out): Intervention: 36/51 (76%) Control (wait list): 35/62 (56%) Subjects who discontinued participation were accounted for at 3 time points and at conclusion of study.	Partly. Mothers interviewed 5 weeks apart at three time points and at 8 month follow-up, a short term period. Longer term effects of intervention not known. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. 'There were no statistically significant differences between the experimental and no treatment comparison groups at baseline on demographic variables and violence severity' (p934). **Was intention to treat (ITT) analysis conducted?** Yes. Using both ITT (include all participants as assigned) and per-protocol analysis (based on data of those who adhered to the treatment protocol). **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. **Were the estimates of effect size given or calculable?** Yes. **Were the analytical methods appropriate?** Yes. Using regression analyses and per-protocol analysis.	Yes. Domestic violence shelter, the research laboratory, or in or near the participant's home. **Does the study relate to at least one of the activities covered by the guideline?** Yes. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Child internalizing problems **Does the study have a UK perspective?** No. USA.	

	Was the precision of intervention effects given or calculable? Were they meaningful? Yes. **Do conclusions match findings?** Yes.		

10. Jouriles EN, McDonald R, Rosenfield D et al. (2010) Improving parenting in families referred for child maltreatment: A randomized controlled trial examining effects of Project Support. Journal of Family Psychology 24: 328–38

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim –The study aimed to evaluate the effectiveness of Project Support, a home-based intervention targeting parenting and maternal distress, in comparison to services as usual, in a sample of families referred for child maltreatment. **Description of theoretical approach?** No. –The study does not report the theory behind the evaluated intervention. **How was selection bias minimised?** Randomised. –Participants were randomly assigned to groups using a random numbers table.	**Was the exposure to the intervention and comparison as intended?** Yes. –A staff person trained in implementing the intervention reviewed audiotapes for 25% of treatment sessions and compared them to therapists' notes to determine how closely the documents reflected the actual content of sessions, indicating 100% correspondence. The review also demonstrated that 52% (SD=20.29) of time in sessions was devoted to the parenting component, with at least 11 of the 12 parenting skills being addressed in 16 of the 17 families assigned to the Project Support condition. Services as usual varied considerably across the 18 families. For example, 4 did not receive any services and of the 14 who did, all received some type of	**Does the study's research question match the review question?** Yes. –In line with the review question, the study aims to evaluate the impact of Project Support in a sample of families reported to CPS for allegations of physical abuse or neglect. **Has the study dealt appropriately with any ethical concerns?** Yes. –All research procedures had Institutional Review Board approval and informed consent was obtained from participating parents. Participants who did not meet eligibility criteria were offered appropriate treatment alternatives. **Were service users involved in the study?**	**Overall assessment of internal validity:** **+** **Overall assessment of external validity:** **++** **Overall validity rating:** **+** Key limitations of the study are: - Assessors not blind to participant condition - Comparison intervention was not consistent across all participants - Method for intent-to-treat analysis not reported

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Was the allocation method followed? Not reported. –The study does not report this. **Is blinding an issue in this study?** No blinding. –The staff person responsible for conducting the baseline assessment was informed of the group assignment, but was blind to the study hypotheses. The same person also conducted follow-up assessments. Mothers were informed of the condition to which they were assigned after the first assessment was completed. **Did participants reflect target group?** Partly. –The sample was small (n= 35), and included a relatively high proportion of families referred for physical abuse rather than neglect, which may reflect a bias. Mothers who were experiencing serious mental health problems or substance abuse disorders were also excluded from the study, whom the authors state can make up a substantial proportion of a	parenting intervention. 12 families also received services in addition to parenting, including anger management, GED classes and individual therapy. **Was contamination acceptably low?** Yes. –Families in the comparison group did not receive the intervention and vice versa. **Did either group receive additional interventions or have services provided in a different manner?** Yes. –12 families also received services in addition to parenting, including anger management, GED classes and individual therapy. **Were outcomes relevant?** Yes. –Reported outcomes clearly relate to the measures used. **Were outcome measures reliable?** Yes. –A variety of measures were used including direct observation of child-parent interactions, data on	No. –Service users involved as participants but not in design or interpretation of results. **Is there a clear focus on the guideline topic?** Yes. –The study is relevant to the guideline topic, focusing on what works to improve outcomes for children and parents. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. –The study population includes mothers of children who are experiencing, or who have experienced, abuse or neglect, as indicated by a substantiated CPS allegation. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. –Project support is a home-based intervention. **Does the study relate to at least 1 of the activities covered by the guideline?**	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
more typical CPS sample. Families were also excluded if they did not speak English. **Were all participants accounted for at study conclusion?** Yes. –There was a relatively low level of attrition. Data were available for 34 of the 35 families at the 8 month assessment (3% attrition) and 31 of the 35 families at the 16 month assessment (11% attrition).	referrals to CPS for child maltreatment and commonly used self-report measures of parenting. The latter included the Parenting Locus of Control Scale (alpha coefficient: .83 at baseline), the Revised Conflict Tactics Scale (alpha coefficient: .82 at baseline) and the Symptom Checklist-90-Revised (alpha coefficient: .98 at baseline). For observed ineffective parenting, Pearson correlations between raters, based on approximately 12 hours of observational data, were .83 for hostile behaviour, .93 for inappropriate use of promotion skills and .76 for inappropriate use of extinguishing skills. **Were all outcome measurements complete?** Yes. –All the planned data was gathered. **Were all important outcomes assessed?** Yes. –The authors report the meaningful effects of the intervention on mothers' parenting versus service as usual, comparing this to other existing studies, in which Project Support exceeds research on par-	Yes. –Intervention aimed primarily at parents and families. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. –Outcomes included mothers' perceived and reported parenting, observed ineffective parenting, maternal psychological distress and recurrence of maltreatment. **(For views questions) Are the views and experiences reported relevant to the guideline?** Not applicable (not views question). **Does the study have a UK perspective?** No. –US study.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	enting programmes with this population. No explicit harms were reported. **Were there similar follow-up times in exposure and comparison groups?** Yes. –Mothers assigned to the Project Support condition were contacted monthly for an additional 8 months (following the 8 month intervention period), and mothers in the comparison condition were contacted monthly for the full 16 months. The authors note that this was so families could be provided with referral information for community resources if desired. **Was follow-up time meaningful?** Yes. –Follow-up is comparable to other similar studies. **Analyses** **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. –There were no statistically significant differences between families assigned to the 2 conditions. The proportion of African American		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	families in the comparison group (61%) exceeded that of the Project Support group (35%), although this difference did not reach statistical significance. **Was intention to treat (ITT) analysis conducted?** Partly. –The authors state that all participants (n=35), including those that were not available at follow-up, were analysed in the groups to which they were originally allocated. However, they do not state how this was conducted, e.g. imputation method for missing data. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Yes. –The authors note that research using similar data concludes that regression coefficients and variance components are estimated without bias, and standard errors of regression coefficients are estimated accurately with a sample size as small as 30 (n=35). Post hoc analysis indicated that the power to detect a medium effect was >.80.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Were the estimates of effect size given or calculable? Partly. –Effect sizes are reported for between-group differences and are calculated using Cohen's standardised effect size. No ES was calculated for the slopes of outcomes over time themselves, as the authors state that there is no agreed general ES statistic for multilevel data. **Were the analytical methods appropriate?** Yes. –The authors used hierarchical linear modeling (HLM), which is appropriate for multilevel data. Separate models were computed for each outcome variable, with outcome scores at each assessment level nested within individuals. Ethnicity was included as a control variable in the analysis. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. –Confidence intervals and p values are reported for the relevant outcome measures. ES for relevant between-groups differences are also reported.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Do conclusions match findings? Yes. –The authors conclude that based on their findings, Project Support may be a promising intervention for improving parenting in families reported to CPS for child maltreatment. It is also noted that, given their small sample size, additional research is needed to further demonstrate the generalisability of these findings. **Were the analytical methods appropriate?** Yes, although methods for intent to treat analysis not reported.		

11. Lieberman AF, Van Horn PJ, Ghosh Ippen C (2005) Toward evidence-based treatment: Child-parent psychotherapy with preschoolers exposed to marital violence. Journal of the American Academy of Child and Adolescent Psychiatry 44: 1241–8

Same trial also reported in: Lieberman AF, Ghosh Ippen C, Van Horn P (2006) Child-parent psychotherapy: 6-month follow-up of a randomized controlled trial. Journal of the American Academy of Child and Adolescent Psychiatry 45: 913–18; Ghosh Ippen C, Harris WW, Van Horn P et al. (2011) Traumatic and stressful events in early childhood: Can treatment help those at highest risk? Child Abuse and Neglect 35: 504–13

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim –The trial sought to evaluate the effectiveness of Child Parent Psychotherapy (CPP) in preschoolers who had been exposed to 'marital	**Was the exposure to the intervention and comparison as intended?** Yes.	**Does the study's research question match the review question?** Yes.	**Overall assessment of internal validity:** **+**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
violence'. The results are reported in 3 papers. One paper reports results up to conclusion of treatment (Lieberman et al. 2005), one at 6 months post-treatment (Lieberman et al. 2006) and one considers effectiveness in relation to exposure to multiple traumatic and stressful events (TSEs) (Ghosh et al. 2011). **Description of theoretical approach?** Yes. –Theoretical premises for the intervention are as follows: - attachment is key organising framework for responses to danger and safety in the first years of life (Ainsworth 1969; Bowlby 1969/1982) - early mental health problems should be addressed in the context of primary relationships (Fraiberg 1980; Lieberman et al. 2000), - child outcomes are a productive of environmental protective and risk factors (Cicchetti and Lynch 1993; Sameroff 1995), - witnessing interpersonal violence can be a traumatic stressor (Pynoos et al. 1999), - the therapeutic relationship is a key factor in treatment (Lieberman et al. 2000) and cultural values of families should be incorporated in to	–Treatment fidelity to CPP was monitored through weekly case supervision, which included review of process notes. The treatment manual was also followed. In the comparison group (individual psychotherapy plus case management), 73% (n=22) of mothers and 55% (n=17) of children received individual treatment, and 45% (n=14) received separate individual psychotherapy for both mother and child. 50% of mothers and 65% of children received more than 20 individual sessions. One child attended fewer than 5 treatment sessions, and 1 mother attended between 5 and 10. The remaining mothers and children attended between 11 and 20 sessions (taken from original study; Lieberman et al., 2005. **Was contamination acceptably low?** Not reported. –This was not reported. **Did either group receive additional interventions or have services provided in a different manner?** No. **Were outcomes relevant?**	–The trial sought to evaluate the effectiveness of Child Parent Psychotherapy (CPP) in preschoolers who had been exposed to 'marital violence'. The results are reported in 3 papers. One paper reports results up to conclusion of treatment (Lieberman et al. 2005), one at 6 months post-treatment (Lieberman et al. 2006) and one considers effectiveness in relation to exposure to multiple traumatic and stressful events (TSEs) (Ghosh et al. 2011). **Has the study dealt appropriately with any ethical concerns?** Yes. –Informed consent was obtained and all procedures received University of California-San Francisco review board approval. **Were service users involved in the study?** No. –Service users were involved as participants, but not in the design of the study or interpretation of results. **Is there a clear focus on the guideline topic?** Yes.	**Overall assessment of external validity:** **+** **Overall validity rating:** **+** Small sample size - Reliance on maternal report - Short follow-up period - In Ghosh (2011), dichotomisation of children into <4 and 4+ TSE risk groups (as nearly all children in the <4 group had experienced at least 2 TSEs).

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
treatment (Tharp 1991; Wessells 1999). The authors also note that marital violence overlaps significantly with child abuse (e.g. Edleson 1999), and that maternal exposure to violence affects quality of parenting. The CPP model is influenced by psychodynamic formulations, social learning and cognitive-behavioural theories and ecological models. **How was selection bias minimised?** Randomised. –Child-parent dyads were randomly assigned to CPP or a comparison group. Method of randomisation not described. **Was the allocation method followed?** Not reported. –This was not reported. **Is blinding an issue in this study?** Part blinding. –Not possible for participants to be blind to treatment condition, but assessors were blind to group assignment. **Did participants reflect target group?**	Yes. –Reported outcomes clearly relate to the measures used. **Were outcome measures reliable?** Partly. –Reliability of all measures not reported. Reliability not reported for measure of Traumatic and Stressful Events (TSEs). The semistructured interview for diagnostic classification DC scored an internal consistency measure of .77 for PTSD and .69 for depression (Kuder-Richardson 20). The Child Behavior Checklist (CBCL; Achenbach 1991, 1992) has been shown to have good reliability, stability and predictive reliability. The Clinician-administered PTSD Scale (CAPS) has excellent test-retest reliability and the Symptoms Checklist-90 Revised (SCL-90-R; Derogatis 1994) has test-retest reliabilities from .78 to .90. The depression scale was used to assess maternal functioning (the reliability of this was not stated). It may also be worth noting that children's outcomes relied on maternal report. **Were all outcome measurements complete?**	–The study covers response to child abuse and neglect. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. –The study population included children aged 3–5, and their mothers, who had been exposed to marital violence, as confirmed by mothers' report on the Conflict Tactics Scale 2 (Straus et al. 1996). **Is the study setting the same as at least 1 of the settings covered by the guideline?** Partly. –Settings are not stated. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. –Study relates to response, intervention is aimed primarily at parents and families. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Yes. –The authors note that minority and low SES children are at greater risk of TSEs, yet are least likely to receive the treatment they need. Participants in this study were therefore, ethnically diverse, with 23% receiving public assistance and 41% having incomes below the federal poverty level (according to the Department of Health and Human Services Guidelines 2004). [Taken from the original studies, Lieberman et al., 2005, 2006.] Mothers were excluded if there was documented abuse of the child, they were currently abusing substances, were homeless, mentally retarded or had psychosis. Children with mental retardation or autistic spectrum disorder were also excluded. **Were all participants accounted for at study conclusion?** Yes. –At post-test, the attrition rate was 14.3% (n=6) in the treatment group and 12% (n=4) in the comparison group. At 6-month follow-up, 2 treatment and 4 comparison dyads dropped from the study, and 7 treatment dyads were not assessed because their treatment	Yes. –All planned data was gathered. **Were all important outcomes assessed?** Partly. –Study examines both child and maternal wellbeing outcomes, however no data on incidence or risk of future child maltreatment was gathered. **Were there similar follow-up times in exposure and comparison groups?** Yes. –Dyads followed up 6 months in to treatment, at conclusion of treatment (reported in Lieberman et al. 2005), 6 months following treatment (reported in Lieberman et al. 2006). **Was follow-up time meaningful?** Partly. –The only longitudinal data collected was at 6 month follow-up, which may impact upon understanding of potential long term benefits. **Analyses**	–Outcomes include both child and maternal functioning. **Does the study have a UK perspective?** No. –US study.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
ended before the 6 month follow-up was added to the study.	**Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. –T tests for continuous variables and Chi-squared tests for categorical variables showed that the groups did not differ on these, as well as demographic variables, dependent variables, or trauma exposure at intake. **Was intention to treat (ITT) analysis conducted?** Yes. –ITT analyses were conducted using a conservative last observation carried forward (LOCF) method, in which the score at the most recent time period was substituted for later incomplete data. Analyses were repeated with the treatment completer (TC) sample, with listwise deletion of cases with missing data to allow for examination of consistency across results. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Partly. –Due to the small sample size, both significant treatment effects and trends in either ITT or TC samples were examined through		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	post-hoc analyses (t tests). The authors did not apply a correction for multiple comparisons due to the risk of increasing a Type II error and focused instead on effect sizes, as has been suggested by other researchers. **Were the estimates of effect size given or calculable?** Yes. –Within group pre to post and pre to follow-up effect sizes were calculated using Cohen's standardised effect size. Chi-squared tests were also used to examine whether pre and posttest groups differed with respect to prevalence of child and maternal PTSD. **Were the analytical methods appropriate?** Yes. –The authors used a general linear model (GLM) repeated measures analysis, which allowed for consistency of analyses across measures, streamlining of data presentation, and decreased post-hoc testing. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	–P values are reported for the relevant outcome measures and the estimates of effect sizes given. **Do conclusions match findings?** Yes.		

12. Linares LO, Li MM, Shrout PE (2012) Child training for physical aggression? Lessons from foster care. Children and Youth Services Review 34: 2416–22

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim –The study aimed to evaluate the effectiveness of an adapted version of the Incredible Years Child Training programme in reducing physical aggression in young children in foster homes. **Description of theoretical approach?** Partly. –The authors do not present a clear theory of change or a logic model but do note that cognitive behavioural approaches such as that used in the intervention are likely to help foster children to develop self-control and coping skills which reduce the risk for physical aggression. **How was selection bias minimised?** Randomised. –Children were randomised at one of the 6 study sites. Method of randomisation is not reported. **Was the allocation method followed?** Not reported.	**Was the exposure to the intervention and comparison as intended?** Not reported. –Information on exposure to the intervention and comparison is not provided. **Was contamination acceptably low?** Yes –The authors do not report on contamination specifically but note that they '… tracked the extent to which knowledge transfer (contamination) may have occurred and found it non-existent' (p2417). **Did either group receive additional interventions or have services provided in a different manner?** No. –There is no indication that either group received additional services or were treated differently by researchers. Differences in levels of mental health service use were not significant at baseline assessment. **Were outcomes relevant?**	**Does the study's research question match the review question?** Yes. –The study aimed to evaluate the effectiveness of an adapted version of the Incredible Years Child Training programme in reducing physical aggression in young children in foster homes. **Has the study dealt appropriately with any ethical concerns?** Yes. –Consent was provided by biological and foster parents however, however approval of the research protocol is not reported. **Were service users involved in the study?** No. –Service users involved as participants only. No indication that service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes.	**Overall assessment of internal validity:** **+** **Overall assessment of external validity:** **++** **Overall validity rating:** **+** A limitation of the study is the very short follow-up period (3 months).

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
–Allocation method and concealment are not reported. **Is blinding an issue in this study?** Part blinding. –Due to the nature of the intervention it would not have been possible to blind participants or providers and it is not clear whether teachers who completed assessments of the child's behaviour and aggression were blinded, however the authors report that interviewers (i.e. those who collected data from foster parents) were blinded to group assignment. **Did participants reflect target group?** Partly. –Quite high numbers of parents declined to participate (31 out of a total eligible sample of 125). It is not clear whether a history of maltreatment was an eligibility criterion and the Guideline Committee may also wish to note that children with a sibling in the study or those whose parents had had their parental rights terminated were ineligible for participation. **Were all participants accounted for at study conclusion?**	Yes. –The study evaluates an intervention designed to improve self-control and reduce physical aggression in foster children and these were measured directly. **Were outcome measures reliable?** Partly. –Outcome measures had established reliability and validity however both were based on parental (biological and foster parents) and teacher reports, particularly given the fact that children may only have been living with their current foster parent for a short period of time. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** No.	–The study evaluates an intervention designed to reduce physical aggression in foster children. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. –The study focuses on physical aggression in foster children all of whom had officially substantiated reports of maltreatment. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. –Response - The study evaluates an intervention which is designed to reduce physical aggression in foster children. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. –Outcomes included physical aggression and self-control.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Yes. –Rates of attrition were acceptable and comparable between groups.	–The final follow-up assessment was conducted at 3 months which is very short and not likely to allow medium term or long term effects to become apparent. The authors do not discuss their rationale for this. **Analyses** **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Partly. –At baseline the authors found significant differences between the 2 groups in terms of gender (greater numbers of male children in the intervention group compared to the usual care group, 59% vs. 38% $\chi 2=3.75$, p=.053), ethnicity) (fewer African American children in the intervention group compared to the usual care group, 37% vs. 62%, $\chi 2=5.39$, p=.020), and Attention Deficit Hyperactivity Disorder diagnosis (greater numbers of children with an initial diagnosis in the intervention group compared to the usual care group, 43% vs. 22%, $\chi 2=4.08$, p=.044). They report that analytic models adjusted for these differences.	**Does the study have a UK perspective?** No. –Study was conducted in the USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Was intention to treat (ITT) analysis conducted? Partly. –The authors report that they took an intent to treat approach but also state that they excluded data from multilevel analyses data from three children who were discharged home during the course of the trial. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** No. –The authors do not present a power calculation or expected effect size, however the sample size is reasonably large (n=94). **Were the estimates of effect size given or calculable?** No. –The authors do not report effect sizes. **Were the analytical methods appropriate?** Yes. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. –P values are provided.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Do conclusions match findings? Yes. –Statistical analysis of teacher ratings is not presented as these showed no change.		

13. Lind T, Bernard K, Ross E et al. (2014) Intervention effects on negative affect of CPS-referred children: Results of a randomized clinical trial. Child Abuse and Neglect 38: 1459–67

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim –The study aimed to determine the effectiveness of Attachment and Biobehavioral Catch-up, an intervention designed to enhance children's self-regulatory capabilities, for young children who had been referred to Child Protective Services. The authors hypothesised that children who participated in the Attachment and Biobehavioral Catch-up programme (with their parents) would display lower levels of negative affect whilst participating in a challenging task than those who participated in the control intervention (with their parents). **Description of theoretical approach?** Yes.	**Was the exposure to the intervention and comparison as intended?** Not reported. –The authors do not provide detail on exposure. **Was contamination acceptably low?** Not reported. –The authors do not provide detail on contamination. **Did either group receive additional interventions or have services provided in a different manner?** No. –There is no indication that either group received additional services. **Were outcomes relevant?**	**Does the study's research question match the review question?** Yes. –The study aimed to determine the effectiveness of Attachment and Biobehavioral Catch-up for young children in families reported to Child Protective Services. They hypothesised that children who received Attachment and Biobehavioral Catch-up '... would show lower expression of negative affect in a challenging task than children in the control intervention group' (p1461). **Has the study dealt appropriately with any ethical concerns?** Yes. –Approval for the study was given by the University of Delaware In-	**Overall assessment of internal validity:** **-** **Overall assessment of external validity:** **++** **Overall validity rating:** **-** The decision to use the Tool Task to observe the child's emotional expression and the use of an unpublished scale to score these (particularly without explanation) and the failure to provide detail on methodological issues such as exposure and contamination mean that it is difficult to be confident in the authors findings.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
–Although the authors do not provide a clear theory of change or logic model they discuss the rationale for the intervention - maltreated children can show emotional development problems and maltreating parents may not provide adequate support to enable their child to develop their emotional regulatory abilities. As Attachment and Biobehavioural Catch-up was designed to encourage parents to respond to their child in a nurturing and non-frightening way, the authors note that they expect the intervention to '... enhance children's developing regulatory capabilities' (p1460). **How was selection bias minimised?** Randomised. –Method of randomisation not reported. **Was the allocation method followed?** Not reported. –Methods of allocation and concealment are not reported. **Is blinding an issue in this study?** Part blinding.	Yes. –The main outcome measured was negative affect although this was measured by combining three subscales of the Revised Manual for Scoring Mother Variables (anger, anger toward parent, global sadness/anger) in the tool-use task which the authors themselves note is a non-clinical measure. **Were outcome measures reliable?** No. –Observations of children's emotion expression were taken using the Tool Task (Matas, L, Arend, RA and Sroufe LA, 1978, Continuity of adaptation in the second year: The relationship between quality of attachment and later competence. Child Development, 49: 547–556) which is a parent-child problem-solving interaction task '... designed to assess children's emotion expression during a challenging task' (p1464). The task was videotaped and the child's behaviour was coded using scales outlined in the 'Revised manual for scoring mother variables in the tool-use task' (Sroufe LA, Matas L, Rosenberg D et al. 1980, University of Minnesota: Unpublished document). The authors	stitutional Review Board and consent was provided by participating families. **Were service users involved in the study?** No. –Service users involved as participants only. No indication that service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. –The study aims to evaluate an intervention designed to enhance self-regulation of emotions in young children whose families had been reported to Child Protective Services. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. –Young children and their families who were involved with Child Protective Services. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
–Due to the nature of the intervention it would not have been possible to blind participants or those delivering the intervention. The authors report that researchers who coded the observational data were blinded. However this is not reported for researchers conducting assessments, for whom this would have been possible. **Did participants reflect target group?** Yes. –Inclusion criteria were: referral to Child Protective Services due to maltreatment allegations, children under the age of two at time of referral, and residence with biological parents. Out of 404 families referred to the programme, only 212 were enrolled and quite significant numbers of those eligible declined to participate (n=32 at the pre-consent stage) or did not respond to contact (n=79 at the pre-consent stage). **Were all participants accounted for at study conclusion?** Yes. –All participants are accounted for. However, there was significant attrition at all stages of the process, including referral to consent	report that inter-rater reliability was acceptable and note that Spearman correlations '… was 0.90 for anger, 0.65 for anger toward caregiver, and 0.62 for global sadness/anger' (p1464). The authors do not discuss their rationale for the use of the Tool Task, an unpublished scoring scale, or the decision to measure negative affect by combining three subscales of this rather than a more established measure. **Were all outcome measurements complete?** Yes. –Although pre-intervention scores could not be collected as the children were too young at this point meaning that changes over time could not be evaluated. **Were all important outcomes assessed?** Yes. –The authors note that the intervention has previously been shown to be effective in relation to cortisol production and attachment levels. **Were there similar follow-up times in exposure and comparison groups?**	–Interventions and assessments were conducted in the family home. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. –Response - prevention/amelioration of impairment. The study focuses on an intervention designed to enhance self-regulation of emotions in young children whose families had been reported to Child Protective Services. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. –Main outcome is negative affect. **Does the study have a UK perspective?** No –Study conducted in USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
(36%), and from consent to enrolment (18%), non-completion of intervention (22% of experimental condition, 18% of control condition).	Partly. –The authors note that the majority of children (62%) participated in the Tool Task assessment over 12 months after the final session of the intervention but it is not clear whether time to follow-up assessment differed by group. **Was follow-up time meaningful?** Partly. –The point at which follow-up assessments were conducted ranged from one month to 27.2 months (M=12.5, SD=6.6) after the final session of the intervention. The mean of 12.5 months would be long enough to detect the effects of the intervention but may not be long enough to assess longer-term benefits or harms. Although the age of the child at enrolment is not provided the mean ages appear appropriate to assess expression and regulation of the child's emotion. **Analyses** **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	–Data is not provided but the authors report that there were no significant differences between groups in relation to: age of child at enrolment, age of child at Tool Task, child's gender, or 'minority status'. There were also no significant differences between groups in terms of age of the parent, their education, or their 'minority status'. **Was intention to treat (ITT) analysis conducted?** Partly. –The authors report that they used an intent-to-treat approach to analysis and included in their analysis all children who '… provided post-intervention data regardless of whether or not the parent completed the intervention' (p1465). However, there was no imputation of missing data from those who completed the intervention but did not provide follow-up data. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. –The authors do not provide a power calculation or expected effect sizes. A sample size of 260 is suitable.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Were the estimates of effect size given or calculable? Yes. –Effect sizes using Cohen's d are provided. **Were the analytical methods appropriate?** Yes. –Analysis of variance with intervention group as the independent variable and the composite measure of negative affect as the dependent variable. Scale scores were transformed into z scores. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. –P values are provided. **Do conclusions match findings?** Yes. **Were the analytical methods appropriate?** Yes.		

14. Mast JE, Antonini TN, Raj SP et al. (2014) Web-based parenting skills to reduce behavior problems following abusive head trauma: A pilot study. Child Abuse and Neglect 38: 1487–95

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim –The study aimed to examine the '... efficacy of a web-based intervention with live coaching designed to improve parenting skills and everyday child functioning ...' (p1488) for children who had experienced abusive head trauma. **Description of theoretical approach?** Partly. –The authors do not provide a logic model or clear theory of change, however they discuss the long-term sequelae of abusive head trauma and note the importance of positive parenting in relation to cognitive development. They also note that specialist care to address the sequelae of abusive head trauma can sometimes be difficult to access and suggest that care provided online has the potential to '... reduce physical and psychological barriers to support and link families with state-of-the-art psychosocial care ...' (p1488). **How was selection bias minimised?** Randomised. –Method not reported.	**Was the exposure to the intervention and comparison as intended?** Not reported. –Information relating to exposure is provided. **Was contamination acceptably low?** Not reported. –Information relating to contamination is provided. **Did either group receive additional interventions or have services provided in a different manner?** No. **Were outcomes relevant?** Yes. –The study focused on parenting skills and child behaviour and the outcome measures used were the Dyadic Parent-Child Interaction Coding Scale, the Eyberg Child Behavior Inventory and the Child Behavior Checklist. **Were outcome measures reliable?** Partly –The measures used have established validity and reliability however data in relation to these are	**Does the study's research question match the review question?** Yes. –The study aimed to examine the '... efficacy of a web-based intervention with live coaching designed to improve parenting skills and everyday child functioning ...' (p1488) for children who had experienced abusive head trauma. **Has the study dealt appropriately with any ethical concerns?** Yes. –The trial was part of a larger study approved by an institutional review board, and families who participated provided consent. **Were service users involved in the study?** No. –Service users involved as participants only. No indication that service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. –The study evaluated an intervention designed to enhance the functioning of children who had suffered abusive head trauma	**Overall assessment of internal validity:** **-** **Overall assessment of external validity:** **++** **Overall validity rating:** **-** Key study limitations: very small sample size (n=9), and resulting very low level of statistical power (12 to 22% at 0.05 criterion - usual standard would be 80%).

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Was the allocation method followed? Not reported. –Details on allocation methods or concealment are not reported. **Is blinding an issue in this study?** Part blinding. –Due to the nature of the intervention it would not have been possible to blind participants or those delivering the intervention. It would have been possible to blind the researchers involved in the collection and coding of data but this is not reported. **Did participants reflect target group?** Partly. –A relatively high number of eligible families declined to participate (5 out of 14 families). **Were all participants accounted for at study conclusion?** Yes. –The overall number of families lost to follow-up was acceptable but this differed by group (intervention n=0, control n=1).	not reported. In addition, the Eyberg Child Behavior Inventory and the Child Behavior Checklist are both based on parental report. In addition, the authors note that coding of parent child interactions which focused on the parent's descriptions of their child's behaviour were removed from analyses due to low inter-rater reliability and that despite close inter-rater reliability for the other categories '... small coding differences greatly affected ICC values for the low frequency categories.' (p 1491) **Were all outcome measurements complete?** Yes **Were all important outcomes assessed?** Yes –Although it is not clear why the authors provide qualitative feedback on the intervention rather than measuring this quantitatively or why this is not contrasted with feedback from parents in the control group. **Were there similar follow-up times in exposure and comparison groups?** Yes	through improving their caregivers parenting skills. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. –Children who have suffered abusive head trauma and their families. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. –The interventions and assessments were delivered in the family home (including the use of videoconferencing). **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. –Response - The study evaluates an intervention designed to improve parenting skills and enhance the functioning of children who have suffered abusive head trauma. **(For effectiveness questions) Are the study outcomes relevant to the guideline?**	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	–At six months after baseline assessment or at programme completion (if sessions had not yet been completed). **Was follow-up time meaningful?** Partly –The six month follow-up assessment would have allowed sufficient time to demonstrate the effects of the intervention on parenting skills and child behaviour however this is unlikely to allow the detection of longer-term beneficial or harmful effects. **Analyses** **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Partly –The authors report that there were no significant differences between groups on outcome measures at baseline (T-test) but the significance of differences between groups on demographic data is not reported. **Was intention to treat (ITT) analysis conducted?** No	Yes –Parent-child interactions and child behaviour. **Does the study have a UK perspective?** No. –The study was conducted in the USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	–The authors report that two families were excluded from analyses (one family from the intervention group dropped out and another from the control group was lost to follow-up). **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Yes –The authors report that 'Power analyses showed that with a sample size of 4 in each group, the ability to detect significant group differences ranged from 12 to 22% power using a .05 significance level' (p 1492). **Were the estimates of effect size given or calculable?** Yes –Relative risk values are provided for the majority of observations coded using the Dyadic Parent-Child Interaction Coding System and partial eta squared effect sizes are provided for scores on the Child Behavior Checklist and the Eyberg Child Behavior Inventory. **Were the analytical methods appropriate?** Yes		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	–The majority of observations using the Dyadic Parent-Child Interaction Coding System were analysed using Poisson regression to compare relative risk. Where data was available from both primary and secondary caregivers this was averaged. Ordinary least squares regression was used to compare child compliance and parental responses to child compliance or non-compliance. ANCOVA was used to analyse parental ratings of child behaviour. The authors report that for measures with low variation across groups, Fisher's exact test was used. Scores were averaged if both caregivers had completed a measure or provided data. The authors did not statistically correct for multiple analysis which they justify on the basis that the study is a pilot study. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes –P values are provided. **Do conclusions match findings?** Partly.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	–The authors appear to discuss data collected during child-directed interactions twice in their narrative and report slightly different p values for these.		

15. Oxford ML, Fleming CB, Nelson EM et al. (2013) Randomized trial of Promoting First Relationships: Effects on maltreated toddlers' separation distress and sleep regulation after reunification. Children and Youth Services Review 35: 1988–92

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim –The study aimed to determine the impact of Promoting First Relationships, an attachment focused intervention, for toddlers recently reunified with their biological parent after being placed in foster care. The study specifically evaluated the impact of the intervention on sleep problems. The authors decided to analyse the subsample of birth parents enrolled in the original study due to their increased risk and significant demographic differences when compared to foster carers or kinship carers participating in the parent study. The authors also note that it is more likely that this group will experience feelings of greater anxiety and inadequacy in the parenting role as a result of the removal of their child from their care	**Was the exposure to the intervention and comparison as intended?** Not reported. –Information on exposure to intervention and comparison is not provided. **Was contamination acceptably low?** Not reported. –Information on contamination is not provided. **Did either group receive additional interventions or have services provided in a different manner?** Partly. –Families in the comparison group were referred to other services when necessary.	**Does the study's research question match the review question?** Yes. –The study aimed to determine the impact of Promoting First Relationships, an attachment focused intervention, for toddlers recently reunified with their biological parent after being placed in foster care. The study specifically evaluated the impact of the intervention on sleep problems and aimed to determine whether this was linked to a reduction in separation distress. **Has the study dealt appropriately with any ethical concerns?** Yes. –Consent was given and the study was approved by an institutional review board.	**Overall assessment of internal validity:** **-** **Overall assessment of external validity:** **++** **Overall validity rating:** **-** Key study limitations: High attrition rate, particularly in intervention group. The exclusion from the study of dyads that were no longer in the same household (presumably those in which the child had been removed back in to care) is a possible source of bias, as these are likely to be families with the highest level of need, for whom the intervention may have been less likely to be effective. Exclu-

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
and would therefore derive greater benefit from the intervention The authors hypothesise that the intervention helps to develop parent's ability to respond sensitively to the behavioural cues of their child and support their ability to self-regulate, thereby leading to reductions in separation distress and sleep problems. **Description of theoretical approach?** Partly. –The authors hypothesise that the intervention would help to develop parent's ability to respond sensitively to the behavioural cues of their child and support the child's ability to self-regulate, thereby leading to reductions in separation distress and sleep problems. **How was selection bias minimised?** Randomised. –Method of randomisation not reported. **Was the allocation method followed?** Not reported. –Allocation methods and concealment are not reported.	**Were outcomes relevant?** Yes. –The study focused on sleep problems and separation distress and these were measured directly. **Were outcome measures reliable?** Yes. –Both measures had established reliability and validity but data to support this are not always presented. The sleep problems measure relied on parental report. **Were all outcome measurements complete?** Yes. –All data was collected as planned. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Partly. –The authors report that due to the fact that the Promoting First Relationships intervention took longer to complete than the comparison intervention there was, on average, more time between	**Were service users involved in the study?** No. –Service users involved as participants only. No indication that service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. –The intervention is designed to improve parenting which is hypothesised to interrupt escalating patterns of separation distress thereby reducing the child's sleep problems. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. –Toddlers who had recently been in foster care and their birth parents. Reasons for placement are not reported, only that placements were all court-ordered (eligibility criterion). **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes	sion of these families may therefore have inflated estimates of the effectiveness of the intervention. The sample size for the study is also relatively small (n=43), and there is no consideration in the paper of study power.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is blinding an issue in this study? Part blinding. –Due to the nature of the intervention it would not have been possible to blind participants or providers, however the authors report that researchers who conducted assessments were blinded. **Did participants reflect target group?** Partly. –The number of eligible participants who agreed to participate is not reported. **Were all participants accounted for at study conclusion?** Partly. –The study states that 'the consenting caregiver and child assessed at baseline, received intervention services, and then assessed post-intervention and 6 months later if they were still in the same household' (p1989). At the 6 month post-intervention assessment 43 dyads remained intact out of a total of 56 who were randomised (25 in the comparison group and 18 in the intervention group). This represents a relatively high overall attrition rate of 23%, and an attrition rate of 33% from within	baseline and 6 month post-intervention assessments for participants in this group than those in the comparison group (10.55 vs 8.80 months, t=3.73, df=41, p<.01). To address this, the researchers included time (in months) between baseline and 6 months post-intervention assessments as a covariate in analyses. **Was follow-up time meaningful?** Partly. –The final assessment took place at 6 months post-intervention which is sufficient to detect the more immediate effects of the intervention but too short to detect more longer-term impacts. **Analyses** **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. –At baseline, children in the intervention group were more likely to have been removed from the care of their birth parent more than once although this was not significant (p=.067 Fischer's exact test). To account for this (and the potential effect multiple removals may have on secure attachment as	–Interventions were delivered, and assessments were conducted in the family home. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. –Response. The intervention is designed to improve separation distress and sleep problems in children who have recently been placed in foster care by enhancing the parenting of their biological parents. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. –Chid sleep problems. **Does the study have a UK perspective?** No. –Study conducted in USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
the intervention group. The exclusion from the study of dyads that were no longer in the same household (presumably those in which the child had been removed back in to care) is a possible source of bias, as these are likely to be families with the highest level of need, for whom the intervention may have been less likely to be effective. Exclusion of these families may therefore have inflated estimates of the effectiveness of the intervention.	well as regulatory or sleep problems) the authors report that '…a dichotomous variable representing multiple removals (yes, no) was included as a covariate in the models' (p1991). **Was intention to treat (ITT) analysis conducted?** No. –Only dyads which remained intact at conclusion were included in analysis. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** No. –Power calculations or expected effect sizes are not presented. The sample is relatively small. **Were the estimates of effect size given or calculable?** Partly. –Effect sizes using Cohen's d are presented in some instances but not consistently. **Were the analytical methods appropriate?** Yes.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Was the precision of intervention effects given or calculable? Were they meaningful? Partly. –P values and confidence intervals are provided inconsistently. **Do conclusions match findings?** Yes.		

16. Purvis KB, Razuri EB, Howard ARH et al. (2015) Decrease in Behavioral Problems and Trauma Symptoms Among At-Risk Adopted Children Following Trauma-Informed Parent Training Intervention. Journal of Child and Adolescent Trauma 8:201–210

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
Study aim Study aim: 'Using a two-group, pre-post intervention design, the current study evaluated the effectiveness of a parent training utilizing Trust-Based Relational Intervention, a trauma-informed, attachment-based intervention, in reducing behavioral problems and trauma symptoms in at-risk adopted children. (p201). **Description of theoretical approach?**	**Was the exposure to the intervention and comparison as intended?** Not reported. There is no discussion of whether the exposure was as intended. **Was contamination acceptably low?** No. It would not have been possible for the control group to receive the intervention. **Did either group receive additional interventions or have services provided in a different manner?** Not reported. It is not discussed whether the intervention group received any additional services, but could affect outcomes. The control group did not receive the on-	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval sought and gained. **Were service users involved in the study?** Yes. Care-givers and children with a history of trauma or 'early adversities'. **Is there a clear focus on the guideline topic?** Partly. It is not 100% clear whether all the participants have experienced abuse	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** +

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
Yes. It is not 100% clear whether all the participants have experienced abuse or neglect. The study states that that they have all experienced 'early adversities'. Participants were asked if they had experienced neglect, physical abuse or sexual abuse. But it is not clear whether the whole sample had experienced one or more. The study uses a theory based three principles that are thought to aid trauma-based care. The intervention is based on the idea that 'felt-safety, self-regulation, and connection' should inform trauma care. The study uses a literature review to make its case that these points are the basis for several interventions that place emphasis on the caregiving relationship. TBRI is an intervention that seeks to aid trauma care by couching it in attachment theory. The	line course until after the intervention group was complete. **Were outcomes relevant?** Yes. The outcomes relate to the measures in the strengths and difficulties questionnaire and the trauma symptoms checklist. **Were outcome measures reliable?** Yes. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Partly. All the outcomes are based on parent feedback. It might have been beneficial to also speak to the child. **Were there similar follow-up times in exposure and comparison groups?** No. **Was follow-up time meaningful?** No. There is a lack of clarity around follow up. Each child was measured against the SDQ and TSCYC before and after the intervention, but there does not appear to have been any follow up beyond that.	or neglect. The study states that that they have all experienced 'early adversities'. Participants were asked if they had experienced neglect, physical abuse or sexual abuse. But it is not clear whether the whole sample had experienced one or more. **Is the study population the same as at least one of the groups covered by the guideline?** Yes. Children and caregivers. **Is the study setting the same as at least one of the settings covered by the guideline?** Yes. Children living with adoptive parents in family homes. **Does the study relate to at least one of the activities covered by the guideline?** Yes. Interventions aimed at supporting children and young people (and their caregivers) who have experienced abuse and neglect. **(For effectiveness questions) Are the study outcomes relevant to the guideline?**	

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
intervention aims to improve caregivers awareness of the child's needs and assist care givers to meet those needs. The intervention seeks to tackle underlying trauma. The theory behind the intervention asserts that by building trusting relationships (or felt-safety) and from here behaviour change (or self-regulation) and connections to others are more possible. **How was selection bias minimised?** Randomised. **Was the allocation method followed?** Partly. Only those that could attend the university on-site sessions were randomised to the intervention group. All the other were randomised to the on-line treatment group or the control group.	**Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. The study states that the groups were matched in terms of age, age at adoption, gender, and type of adoption (international or domestic). **Was intention to treat (ITT) analysis conducted?** Not reported. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. **Were the estimates of effect size given or calculable?** Yes. **Were the analytical methods appropriate?** Yes. The two subscales were examined by repeated measures Multivariate Analysis of Covariance (MANCOVAs). **Do conclusions match findings?** Yes.	Yes. The outcomes fall under 'children and young people's health and wellbeing and quality of parenting and parent-child relationships. **Does the study have a UK perspective?** No. USA.	

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
Is blinding an issue in this study? Blinding not possible **Did participants reflect target group?** Partly. The sample were chosen, in the first instance on their ability to attend the on-site training course, they were recruited via advertisement. It is hard to say whether the backgrounds of the children were assessed before recruitment. The sample was uneven in terms of gender, ethnicity, adoption type (domestic or international) and type of abuse. However, the control group were matched to the intervention group on age, gender. adoption type and age at adoption. **Were all participants accounted for at study conclusion?** Yes.			

17. Reddy SD, Negi LT, Dodson-Lavelle B et al. (2013) Cognitive-Based Compassion Training: A promising prevention strategy for at-risk adolescents. Journal of Child and Family Studies 22: 219–30

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim –Aim of study to examine whether a 6-week Cognitive-Based Compassion Training (CBCT) intervention would improve psychosocial functioning among adolescents in foster care. **Description of theoretical approach?** Yes. –The study states that 'Building on basic mindfulness practice, Cognitively-Based Compassion Training (CBCT) employs a variety of cognitive restructuring and affect generating practices with the long-term goal of developing an equanimity of mind that fosters acceptance and understanding of others (Salzberg 2002)' (p220). **How was selection bias minimised?** Randomised. –Block randomisation. **Was the allocation method followed?** Yes.	**Was the exposure to the intervention and comparison as intended?** Not reported. **Was contamination acceptably low?** Not reported. **Did either group receive additional interventions or have services provided in a different manner?** Not reported. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Partly. –Study used standardised measures, however one measure (Self-Other Four Immeasurables Scale, Kraus and Sears 2009) has not been used with adolescents, although reliability was acceptable(Cronbach's alpha = 0.63). **Were all outcome measurements complete?** No.	**Does the study's research question match the review question?** Yes. –Study is examining the impact of an intervention aimed at improving psychosocial functioning among adolescents in foster care. **Has the study dealt appropriately with any ethical concerns?** Partly. –All procedures were approved by the Georgia Department of Human Services Internal Review Board. Participants gave informed consent. Waitlist control. However, study does not seem to have considered ethical implications of teaching maltreated children compassion towards 'strangers and enemies', given what is known about tendency amongst some maltreated children and young people to be vulnerable to exploitation by others. **Were service users involved in the study?** No.	**Overall assessment of internal validity:** **-** **Overall assessment of external validity:** **+** Unable to award ++ due to lack of specificity regarding maltreatment history of participants. **Overall validity rating:** **-** Poor reporting of sample size and attrition rates. Unclear whether assessors were blinded to treatment condition. No consideration of statistical power. Short follow-up time.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is blinding an issue in this study? No blinding. –It was not possible to blind participants or (presumably) their legal guardians (who completed one of the outcome measures) to intervention condition. It is not stated whether the researchers administering the various outcome scales were blind to intervention condition or not. **Did participants reflect target group?** Yes. –Participants were young people aged 13–17 in foster care. However, participants were 'free of active suicidality, psychotic disorders, bipolar I disorder, eating disorders and chronic illness' (p220). This could mean that the sample population has a lower level need than the foster care population as a whole. **Were all participants accounted for at study conclusion?** Not reported. –Retention rates not reported.	–Study states that 'due to a research staff error, caregivers of wait-list participants were only given the ICU-p at baseline but not Study Week 6' (p223). Statistical analyses do not appear to have been conducted in relation to all variables. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Partly. –Follow-up was conducted at 6 weeks only (after completion of treatment). This is a relatively short timeframe within which to observe improvements, and also provides no information about longer term effects of the intervention. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. –Used independent samples t tests to confirm that there were no	–Service users as participants only - not involved in design or interpretation of results. **Is there a clear focus on the guideline topic?** Yes. –Study focuses improving psychosocial functioning among adolescents in foster care, stating that 'these youth have suffered maltreatment'. **Is the study population the same as at least one of the groups covered by the guideline?** Partly. –Study population is young people in foster care. These young people were selected due to the 'exceptionally high rates of maltreatment' in this group (p219). However, it is not documented in the study whether all participants have experienced abuse or neglect. **Is the study setting the same as at least one of the settings covered by the guideline?** Yes. **Does the study relate to at least one of the activities covered by the guideline?**	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	demographic differences by group, or differences in terms of mood, behaviour and emotion regulation at baseline. **Was intention to treat (ITT) analysis conducted?** Not reported. –Unclear - and unclear whether there was attrition from the study. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. **Were the estimates of effect size given or calculable?** Partly. –Effect sizes not reported, but are calculable from available data. **Were the analytical methods appropriate?** Partly. –Analysis of covariance - appropriate as existing correlations between sample characteristics and outcome measures. Analyses controlled for age, ethnicity, gender and baseline scores. **Do conclusions match findings?** Yes.	Yes. –Study relates to response following abuse or neglect. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. –Outcome measures explore young people's wellbeing. **Does the study have a UK perspective?** No. –US study. However, service delivery context is likely to be similar to this country, that is, a specialist intervention delivered to young people in foster care.	

18. Rushton A, Monck E, Leese M et al. (2010) Enhancing adoptive parenting: A randomized controlled trial. Clinical Child Psychology and Psychiatry 15: 529–42

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim –The study aimed to evaluate 2 parenting programmes designed to improve adoptive parenting and child behavioural problems. **Description of theoretical approach?** No. –The authors do not discuss the theory underlying the intervention. **How was selection bias minimised?** Randomised. –Permuted block randomisation. **Was the allocation method followed?** Not reported. –Allocation methods and concealment are not reported. **Is blinding an issue in this study?** Blinding not possible. –Due to the nature of the intervention it would not have been possible to blind participants or providers and the authors note that as follow-up interviews included a focus on the adopters' involvement	**Was the exposure to the intervention and comparison as intended?** Not reported. –Information on exposure to the interventions is not provided. **Was contamination acceptably low?** Not reported. –Information on contamination is not provided. **Did either group receive additional interventions or have services provided in a different manner?** No. –Although the authors report that some participants in the control group received support (unspecified content) they state that this was '... far less intensive than the individualized parenting advice provided in the trial' (p532). **Were outcomes relevant?** Yes. –The study focused on adoptive parenting and child behaviour problems and the outcome measures used included the	**Does the study's research question match the review question?** Yes. –The study aimed to evaluate '... two parenting programmes designed for adopters of children late placed from care' (p529). **Has the study dealt appropriately with any ethical concerns?** Yes. –Consent was provided by adoptive parents and the study was approved by a research ethics committee. **Were service users involved in the study?** No. –Service users involved as participants only, no indication of involvement at design stage or in interpretation of findings. **Is there a clear focus on the guideline topic?** Yes. –The study evaluated two parenting programmes aiming to improve the adoptive parenting of	**Overall assessment of internal validity:** **-** **Overall assessment of external validity:** **++** **Overall validity rating:** **-** The small sample size and use of a scale with unclear reliability or validity, as well as a lack of detail on key methodological issues such as the use of intent to treat analysis means that it is not possible to award a higher score. The analysis also combines the 2 parenting interventions, making it difficult to draw conclusions about what has led to any improvement in outcomes.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
in either condition it was not possible for these to be blinded. **Did participants reflect target group?** Partly. –Out of 80 families screened as eligible only 38 participated in the study. One family withdrew from the study before baseline assessments were conducted but it is not clear if this occurred before randomisation. Children who had been placed with a relative or with their existing foster parents were ineligible. **Were all participants accounted for at study conclusion?** Yes. –Drop-out rates were acceptable.	Strengths and Difficulties Questionnaire, the Expression of Feelings Questionnaire, the Post Placement Problems scales, Parenting Sense of Competence Scale, and the Daily Hassles scale. **Were outcome measures reliable?** Partly. –The majority of outcome measures had established reliability and validity however they all rely on self-report and no data in relation to reliability and validity is provided. In addition, the Post Placement Problems scale was created specifically for use in this study in order to measure the post-adoption experiences of maltreated children, however the internal consistency of this scale is not reported. The authors also report that they measure parental management of emotional difficulties and provide statistical analysis of this but this appears to be based on qualitative data with no explanation of how this was collected (e.g. what questions were used), as a result only the qualitative findings, rather than quantitative calculations, from this section have been extracted.	children late placed from care with serious behavioural problems. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. –Late placed adoptive children between the ages of 3 and 8 with serious behavioural problems. The majority had experienced some form of abuse or neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. –Interviews and sessions were conducted and delivered in the family home. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. –The study evaluates 2 parenting programmes which aimed to improve adoptive parenting and reduce behavioural difficulties in children late placed from care. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Were all outcome measurements complete? Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. –Both groups were followed up for the same length of time (6 months post-intervention). **Was follow-up time meaningful?** Yes –The final follow-up interview was conducted at 6 months post-intervention which would have been sufficient to detect the impact of the intervention on parenting but may not be long enough to allow the longer-term effects on child behaviour to become apparent. **Analyses** **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes.	–Outcomes included children's psycho-social problems; nature of the relationship between the child and their adoptive parent; post-adoption problems; and perceived parenting competence and challenges. **Does the study have a UK perspective?** Yes. –Study was conducted in England.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	–The authors report that the groups were '... well balanced ...' (p535) but no statistical tests were conducted to assess differences between the groups. **Was intention to treat (ITT) analysis conducted?** Not reported. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** No. –The authors report that in order to detect a difference in power at 0.8 ($p<.05$) they required a sample of 27 families in each group, when in fact the sample sizes were 19 and 18 respectively. **Were the estimates of effect size given or calculable?** Yes. –Effect sizes using Cohen's are reported. **Were the analytical methods appropriate?** Yes. –Analysis of variance (controlled for baseline variables), linear regression and Chi-squared tests. NB Due to the small sample size		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	the authors report that they combined the two experimental groups in their analysis. **Was the precision of intervention effects given or calculable? Were they meaningful?** Partly. –P values and confidence intervals (for some outcome measures) are provided. **Do conclusions match findings?** Partly. –In their discussion the authors emphasise the effect which the interventions had on parenting sense of competence which does not directly correspond to their aim of improving parenting.		

19. Spieker SJ, Oxford ML, Kelly JF et al. (2012) Promoting First Relationships: Randomized trial of a relationship-based intervention for toddlers in child welfare. Child Maltreatment 17: 271–86

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim –The study aimed to evaluate the impact of the Promoting First Relationships intervention for caregivers of toddlers with a recent placement in foster care. The au-	**Was the exposure to the intervention and comparison as intended?** Not reported. –Information on exposure to intervention and comparison is not provided.	**Does the study's research question match the review question?** Yes –The study aimed to evaluate the impact of the Promoting First Re-	**Overall assessment of internal validity:** **-** **Overall assessment of external validity:** **++**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
thors hypothesised that the intervention '… would result in improved parenting and child outcomes relative to a comparison condition in which families received home-based services that were not relationship focused' (p273) **Description of theoretical approach?** Partly. –The authors hypothesise that the intervention would improve caregiver's abilities to recognise which child behaviours correspond to a need for nurturance (e.g. unmet emotional needs displayed through difficult behaviours) and that this would enhance parental sensitivity. This in turn is expected to promote more secure attachments and improved regulation of emotions in the child. **How was selection bias minimised?** Randomised. –Computer generated and blocked by caregiver type. **Was the allocation method followed?** Not reported.	**Was contamination acceptably low?** Yes. –The authors report that fidelity of the comparison intervention was monitored and that this indicated that no intervention strategies were used in these sessions. **Did either group receive additional interventions or have services provided in a different manner?** Partly. –The authors report that providers delivering the comparison intervention helped participants in this group to access other services such as Early Head Start, mental health services housing, etc. **Were outcomes relevant?** Yes. –The authors hypothesised that the intervention would improve parenting, child attachment and emotional regulation; and the measures used were relevant to these outcomes. **Were outcome measures reliable?** Yes.	lationships intervention for caregivers of toddlers with a recent placement in foster care. **Has the study dealt appropriately with any ethical concerns?** Yes. –Informed consent and approval of the study by an institutional review board or ethics committee are not reported specifically but the authors note that recruitment for the study involved both an institutional review board, and state social services. **Were service users involved in the study?** No. –Service users involved as participants only. No indication that service users were involved at the design stage. **Is there a clear focus on the guideline topic?** Yes. –The intervention is designed to improve parenting (e.g. greater sensitivity) which is hypothesised to promote the development of secure attachment and improved emotional regulation in toddlers who have recently been placed in foster care.	**Overall validity rating:** **+** Key study limitations: relatively high attrition rate from the study. Relatively high numbers of participants did not complete follow up assessments (n=28 (26%) at post-intervention stage, and n=34 (32%) at the 6 month point.) At the immediate post-intervention assessment this was comparable by group but at the six month point greater numbers of dyads in the intervention group (n=22) failed to complete assessments than those in the control group (n=12). In addition, significant numbers of dyads were excluded from the analyses due to changes in caregiver throughout the course of the study. This meant that data from only 56% of participants in intervention group, and 66% of comparison group were used in the analysis.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
–Details on allocation methods and concealment are not provided. **Is blinding an issue in this study?** Part blinding. –Due to the nature of the intervention it would not have been possible to blind participants or providers to group assignment, however the authors do report that researchers who conducted assessments were blinded. **Did participants reflect target group?** Partly. –A relatively high number of eligible participants declined to take part (61 out of 271 determined to be eligible). **Were all participants accounted for at study conclusion?** Partly. –Relatively high numbers of participants did not complete follow up assessments (n=28 (26%) at post-intervention stage, and n=34 (32%) at the six month point.) At the immediate post-intervention assessment this was comparable by group but at the 6 month point greater numbers of dyads in the	–All outcomes appear to have established reliability and validity although data to support this is not always presented. The majority of measures relied on self-report data. **Were all outcome measurements complete?** No. –The authors report that '… some measures had further missing data (<5%) due to observational data being uncodable.' The measures affected by this are not reported. Child internalising problems, externalising problems, sleep problems and 'other problems' (measured using subscales of the Child Behavior Checklist for Ages 1½–5) were not measured at baseline or the immediate post-intervention assessment due to the young age of the children. Assessments of child emotional regulation and orientation (measured using scales from the scales Bayley Behavior Rating Scales) were not conducted at the immediate post-intervention assessment due to concerns that the interval between completion of the intervention and the assessment was too short.	**Is the study population the same as at least 1 of the groups covered by the guideline?** Yes –Caregivers of toddlers who had recently been in foster care (between the ages of ten and 24 months). Reasons for placement are not reported, only that placements were all court-ordered (eligibility criterion). **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. –Assessments and intervention sessions were conducted in the caregiver's home. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. –Response. The intervention is designed to improve the attachment levels and emotional regulation of children who have recently been placed in foster care by enhancing the parenting of their caregivers.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
intervention group (n=22) failed to complete assessments than those in the control group (n=12). In addition, significant numbers of dyads were excluded from the analyses due to changes in caregiver throughout the course of the study. This meant that data from only 56% of participants in intervention group, and 66% of comparison group were used in the analysis.	**Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Partly. –The final assessment took place at 6 months post-intervention which is sufficient to detect more immediate effects of the intervention but too short to detect more longer-term impacts. **Analyses** **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Partly –At baseline, children in the intervention group were significantly more likely to have been removed from the care of their birth parent more than once $\chi 2$ (1, n=210) = 7.31, p<.01. The ANCOVA model adjusted for this difference. **Was intention to treat (ITT) analysis conducted?** Partly.	**(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. –Outcomes include parental sensitivity and stress; as well as the child's attachment levels, behaviour and emotional regulation. **Does the study have a UK perspective?** No. –Study conducted in USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	–The authors report that their analysis was carried out using intention to treat models however they also state that their analysis only included dyads which remained intact throughout the course of the study. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Yes. –The number of dyads which remained intact at the conclusion of the study was sufficient to yield power of .80 to detect an effect size of d=.50 with an α of p<.05. **Were the estimates of effect size given or calculable?** Yes. –Effect sizes using Cohen's d are provided. **Were the analytical methods appropriate?** Yes. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. –P values are provided.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Do conclusions match findings? Yes. –Although discussion really only focuses on outcome measures for which significant differences were detected at either follow-up assessment.		

20. Stronach EP, Toth SL, Rogosch F et al. (2013) Preventive interventions and sustained attachment security in maltreated children. Development and Psychopathology 25: 919–30

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim –The aim of the study was to evaluate the efficacy of a child–parent psychotherapy programme and a psychoeducational parenting intervention in comparison to care as usual. The study focused on secure attachment and behavioural functioning at 12 months (expanding on the findings from a previous study in which assessments were conducted in the immediate post-intervention period - Cicchetti D, Rogosch F, Toth SL, 2006, Fostering secure attachment in infants in maltreating families through preventive interventions. Development and Psychopathology, 18: 623–49). NB The NCCSC have reported the findings of this paper as part of the	**Was the exposure to the intervention and comparison as intended?** Not reported. –The authors do not provide details on exposure to the intervention. **Was contamination acceptably low?** Not reported. –The authors do not provide details on contamination. **Did either group receive additional interventions or have services provided in a different manner?** Not reported.	**Does the study's research question match the review question?** Yes. –The study aimed to expand on the findings of a previous paper which evaluated the relative efficacy of child–parent psychotherapy and a psychoeducational parenting intervention in comparison to standard care. Data from 12 months post-intervention was analysed. (Cicchetti D, Rogosch F, Toth SL, 2006, Fostering secure attachment in infants in maltreating families through preventive interventions. Development and Psychopathology, 18: 623–49). **Has the study dealt appropriately with any ethical concerns?**	**Overall assessment of internal validity:** **+** **Overall assessment of external validity:** **++** **Overall validity rating:** **+**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
findings extracted from Goldman Fraser J, Lloyd SW, Murphy RA et al. (2013) Child exposure to trauma: Comparative effectiveness of interventions addressing maltreatment. Comparative Effectiveness Review, 89: 1–161. **Description of theoretical approach?** No. –The authors do not clearly outline the theories underlying the interventions although they do discuss potential effects which they may have on parenting. **How was selection bias minimised?** Randomised. –Method of randomisation not reported. **Was the allocation method followed?** Not reported. –Allocation methods and concealment are not reported. **Is blinding an issue in this study?** Part blinding. –Due to the nature of the intervention it would not have been possi-	–It is unclear whether children in the non-maltreated comparison group received any services. **Were outcomes relevant?** Yes. –The study focused on secure attachment and behavioural functioning and these were measured directly using appropriate scales. **Were outcome measures reliable?** Partly. –The scales used had established reliability and validity however scores on the Child Behavior Checklist were based on maternal reports of child behaviour. The authors note that because abusive parents may overemphasise externalising behaviours, scores on this scale were '... interpreted in terms of parental perceptions of behavior problems' (p924). In addition, children's attachment levels appear to have been coded and classified using different manuals at baseline, immediate post-intervention and 12 month post-intervention follow-up assessments using different manuals and this is not explained by the authors.	Partly. –Informed consent processes are not reported explicitly but the authors note that participating families were made aware that involvement was voluntary and a decision not to participate would not affect their receipt of any other services. Approval of study protocol is not reported. **Were service users involved in the study?** No. –Service users involved as participants only, no indication of involvement at design stage or interpretation of findings. **Is there a clear focus on the guideline topic?** Yes. –The study evaluates the effects of two interventions on attachment in maltreated children. **Is the study population the same as at least one of the groups covered by the guideline?** Yes. –The majority of participants were maltreated children and their families.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
ble to blind participants or providers to group allocation. The authors report that researchers who coded videotaped 'Strange Situation' assessments for children's attachment levels were blinded to group assignment but blinding of other investigators is not reported. **Did participants reflect target group?** Partly. –The number of families screened as eligible and those who agreed to participate are not reported. **Were all participants accounted for at study conclusion?** Yes. –All participants were accounted for however attrition rates were relatively high. The authors report that 32 families (60.4%) randomised to the child–parent psychotherapy group and 24 families (48.9%) randomised to the psychoeducational parenting intervention '... participated in the interventions' (p924); 41 families (21.7%) did not complete postintervention assessments (12 months). The authors note that the percentage of families unavailable at this stage was significantly higher in the care as usual group (33.4%,	**Were all outcome measurements complete?** Yes. –Maternal perceptions of child behaviour only assessed at 12 months post-intervention follow-up. **Were all important outcomes assessed?** Yes. –Although the Guideline Committee may wish to note that attachment was assessed during the 'Strange Situation' procedure in a research environment. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Yes. –The study analyses data collected at the 12 month point which may not have been sufficient to detect longer-term benefits or harms. **Analyses** **Were exposure and comparison groups similar at baseline? If not, were these adjusted?**	**Is the study setting the same as at least one of the settings covered by the guideline?** Yes. –Interventions and some assessments took place in the family home. **Does the study relate to at least one of the activities covered by the guideline?** Yes. –Response (prevention of impairment) - The study evaluates the effects of 2 interventions on attachment in maltreated children. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. –The study focused on children's attachment and behavioural problems. **Does the study have a UK perspective?** No. –The study was conducted in the USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
n=27) than in other groups. The number of families available at follow-up assessment is not reported specifically but the data provided in the tables suggests n=145.	Yes. **Was intention to treat (ITT) analysis conducted?** Partly. –The intent-to-treat analysis does not include all comparisons used in treatment completer analyses. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. –Power calculations or expected effect size are not presented however the sample size seems reasonable. **Were the estimates of effect size given or calculable?** Partly. –Effect sizes are presented but it is not clear what type has been used and it is therefore not possible to include a narrative description of the size of effect when reporting the authors findings. **Were the analytical methods appropriate?** Yes. **Was the precision of intervention effects given or calculable? Were they meaningful?**		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Yes. –P values are provided. **Do conclusions match findings?** Yes.		

21. Swenson CC, Schaeffer CM, Henggeler SW et al. (2010) Multisystemic therapy for child abuse and neglect: A randomized effectiveness trial. Journal of Family Psychology 24: 497–507

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim –To evaluate an adaptation of multisystemic therapy for physically abused adolescents and their families. **Description of theoretical approach?** Yes. –based in part on ecological systems theory. **How was selection bias minimised?** Randomised. –Randomisation using a computer-generated table of random numbers. **Was the allocation method followed?** Yes.	**Was the exposure to the intervention and comparison as intended?** Yes. **Was contamination acceptably low?** Not reported. **Did either group receive additional interventions or have services provided in a different manner?** Partly. –MST sessions titrated according to family needs (see intervention details), Control group (EOT) also had additional services provided when required (see comparison details). **Were outcomes relevant?**	**Does the study's research question match the review question?** Yes. –Study is examining effectiveness of therapeutic intervention following physical abuse. **Has the study dealt appropriately with any ethical concerns?** Yes. –Written informed consent obtained; approved by the institutional review board of the participating university. If family chose not to participate, the caseworker arranged other treatments. Youth were compensated $15 and parents $35 per assessment. **Were service users involved in the study?**	**Overall assessment of external validity:** + **Overall assessment of internal validity:** + **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
–Allocation by opening a sealed envelope and family informed of the assigned treatment condition. **Is blinding an issue in this study?** No blinding. **Did participants reflect target group?** Yes. –Some youth were in placement. **Were all participants accounted for at study conclusion?** Yes. –Intention-to treat analysis, 44/45 participants in intervention and 42/45 participants completed study. Retention was 100% through months 2 and 4, and 97% through months 10 and 16.	Yes. **Were outcome measures reliable?** Yes. –All validated. **Were all outcome measurements complete?** Partly. –Service use not reported. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. –At 2, 4, 10 and 16 months. **Was follow-up time meaningful?** Yes. **Analyses** **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** **Was intention to treat (ITT) analysis conducted?** –Yes.	No. –Service users involved as participants only - not in design or interpretation of results. **Is there a clear focus on the guideline topic?** Yes. –Study is relevant to response following abuse and neglect. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. –Population is young people and custodial parents known to Child Protective Services due to physical abuse. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. –Community setting. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. –Study is relevant to response. **(For effectiveness questions) Are the study outcomes relevant to the guideline?**	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Were exposure and comparison groups similar at baseline? If not, were these adjusted? Yes. –Comparisons conducted using chi-square and t tests, and no significant differences between baseline and comparison group. **Was intention to treat (ITT) analysis conducted?** Yes. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Yes. –Power calculation: limited for detecting medium effects; but adequate for large effects (p502). **Were the estimates of effect size given or calculable?** Yes. –Effect sizes reported using Cohen's d. **Were the analytical methods appropriate?** Yes. –Latent growth curve modelling. **Was the precision of intervention effects given or calculable? Were they meaningful?**	Yes. **Does the study have a UK perspective?** No. –US study. Service context differs from UK, particularly in terms of comparison intervention which include 'standard services provided for physically abused youths and their parents'. There is insufficient description of 'standard services' to know whether this is similar to UK standard services.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Yes. –P values, effect sizes and standard error given. **Do conclusions match findings?** Yes.		

22. Toth SL, Sturge-Apple ML, Rogosch FA et al. (2015) Mechanisms of change: Testing how preventative interventions impact psychological and physiological stress functioning in mothers in neglectful families. Development and psychopathology 27: 1661–74

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim To identify the impact of two preventative interventions - Child-Parent Psychotherapy(CPP) and Pyschoeducational Parenting Intervention(PPI) on levels of maternal stress in mothers from neglectful families, compared with community standard treatment for maltreating parents (CS) and a nonmaltreating comparison group (NC). **Description of theoretical approach?** Yes. Study based on theory	**Was the exposure to the intervention and comparison as intended?** Yes. **Was contamination acceptably low?** Not reported. **Did either group receive additional interventions or have services provided in a different manner?** Not reported. **Were outcomes relevant?** Partly. Parental stress for proxy as predictor of future maltreatment. **Were outcome measures reliable?** Partly. Parenting stress questionnaire is a val-	**Does the study's research question match the review question?** Partly. Levels of parenting stress used as a proxy for risk of maltreatment. **Has the study dealt appropriately with any ethical concerns?** Partly. States that mothers gave written consent and several screening levels for nonmaltreating mothers. No ethical statement on authors' manuscript reviewed. **Were service users involved in the study?** No **Is there a clear focus on the guideline topic?** Yes. Study discusses link between parenting stress and maltreatment of children.	**Overall assessment of internal validity:** + More precision on results in providing confidence intervals for differences between groups and a power calculation combined with no blinding. **Overall assessment of external validity:** + Lack of UK focus and doesn't exactly address review question. **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
that maltreating parents experience higher levels of parenting stress (cites Haskett et al.2006, McCanne & Hagstrom 1996 on p.5) while identifying a research gap of understanding impact of pyschosocial intervention on biological processes (p.5). In terms of physical hypothalamic-pituitary-adrenal(HPA)-Axis functioning as measure of biological stress cites family risk models which suggest that physiological responses to family stressors serve as an explanatory mechanism in links between family adversity and functioning (p.6, cites Repetti, Taylor & Seeman 2002). The authors have found that dysregulation in basal activity in the HPA axis has been linked to 'perturbations in caregiving' in their previous work (p.6). Logic models for	idated tool and measuring basal cortisol activity accepted method of measuring physiological stress but link/logic model new to this study. **Were all outcome measurements complete?** Not reported. **Were all important outcomes assessed?** Partly. Could have been validated by looking at CPS reports. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Yes. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. Differences in levels of psychological stress in CPP group at baseline controlled for. **Was intention to treat (ITT) analysis conducted?** Yes.	**Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Mothers who maltreat their children. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Home of children/biological mothers. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to Response. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **Does the study have a UK perspective?** No, US study.	More precision on results in providing confidence intervals for differences between groups and a power calculation combined with no blinding

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
how psychological and physiological stress factors are associated are presented in Figures 1 and 2 in pp.22-23. **How was selection bias minimised?** Randomised Matched groups **Was the allocation method followed?** Yes **Is blinding an issue in this study?** Blinding not possible **Did participants reflect target group?** Yes **Were all participants accounted for at study conclusion?** Not reported	**Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. **Were the estimates of effect size given or calculable?** Partly. Does not give confidence intervals and only describes mean difference in text not included in data table. **Were the analytical methods appropriate?** Yes. **Was the precision of intervention effects given or calculable? Were they meaningful?** No. Confidence intervals and power calculation not stated. **Do conclusions match findings?** Yes.		

23. Winokur M, Ellis R, Drury I et al. (2014) Answering the big questions about differential response in Colorado: Safety and cost outcomes from a randomized controlled trial. Child Abuse and Neglect 39: 98–108

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim 1) To assess the impact on child safety outcomes of a family assessment response versus an investigation response assigned to children and families with a referral for child neglect or abuse 2) To examine the cost implications for child welfare agencies that implement a DR (Differential Response)-organized CPS (child protective services) system. **Description of theoretical approach?** No. **How was selection bias minimised?** Randomised. **Was the allocation method followed?** Not reported. Method of randomisation, allocation concealment not reported.	**Was the exposure to the intervention and comparison as intended?** Yes. **Was contamination acceptably low?** Not reported. **Did either group receive additional interventions or have services provided in a different manner?** Not reported. **Were outcomes relevant?** Yes. Child welfare outcomes such as referrals, assessments, high risk assessments recorded later than 3 days after the initial referral. **Were outcome measures reliable?** Yes. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Not reported.	**Does the study's research question match the review question?** Yes. To assess the effect/impact of differential response to reports of child abuse. **Has the study dealt appropriately with any ethical concerns?** No. Not reported. **Were service users involved in the study?** Yes. Service users were participants in the trial. **Is there a clear focus on the guideline topic?** Yes. To compare the effect/impact of differential response to reports of child abuse. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Families with a referral for child neglect or abuse. **Is the study setting the same as at least 1 of the settings covered by the guideline?** No. Not reported. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Response to child neglect and abuse.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is blinding an issue in this study? No blinding. Caseworkers were not randomly assigned to serve FAR or IR cases. **Did participants reflect target group?** Yes. Families with a referral for child neglect or abuse. **Were all participants accounted for at study conclusion?** Not reported. attrition or drop-out rate not reported.	**Was follow-up time meaningful?** Not reported. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Partly. No statistically significant differences between FAR and IR families except: 1. IR families had more children in the home than did FAR 2. FAR families had an 'older' youngest child than did IR families 3. IR families had more caregivers than did FAR families. These characteristics were adjusted in the regression models. **Was intention to treat (ITT) analysis conducted?** Yes. Used an 'intent-to-treat' (ITT) analysis. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. **Were the estimates of effect size given or calculable?** Yes. Effect size given. **Were the analytical methods appropriate?**	**(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **Does the study have a UK perspective?** No, Colorado, USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Yes. Regression and survival analyses on safety outcomes. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. Effects given. **Do conclusions match findings?** Yes.		

Review question 15 – Findings tables

1. Barlow J, Johnston I, Kendrick D et al. (2006) Individual and group-based parenting programmes for the treatment of physical child abuse and neglect. Cochrane Database of Systematic Reviews issue 3: CD005463

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim – To assess the efficacy of brief (i.e. between 6 and 30 weeks) group-based or 1:1 parenting programmes in addressing child physical abuse or neglect. **Methodology** Systematic review. **Country**	**Participants** Caregivers and families – Studies were eligible for inclusion in the review if the intervention was provided directly to parents of children aged 0–19 years. Programmes had to have targeted parents who have a history of physical abuse or neglect (p5). **Sample characteristics** Age – not reported. Sex	Effect sizes A. Parenting programs vs. control (no active treatment): child abuse potential: 1). a large significant difference favouring the intervention group SMD (Standard Mean Difference) -0.99 [-1.71 to -0.27] (Terao 1999). Parental involvement - free play: 2). no difference SMD-0.76 [-1.56, 0.04] (Hughes 2004). Parental involvement - ring toss 3). no difference SMD-0.34 [-1.12, 0.43] (Hughes 2004).	**Overall assessment of internal validity** + **Overall assessment of external validity** + **Overall score** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Range of countries. **Source of funding** Voluntary/Charity – the Nuffield Foundation UK	– parents Ethnicity – not reported. Religion/belief – not reported. Disability – not reported. Long term health condition – not reported. Sexual orientation – not reported. Socioeconomic position – not reported. Type of abuse – physical abuse (5 RCTs); physical abuse and neglect (1 RCT); unspecified type of abuse (1 RCT). Looked after or adopted status – not reported. Unaccompanied asylum seeking, refugee or trafficked children – not reported. **Sample size** Systematic reviews: number of studies – 7 RCTs published between 1983–2004. **Intervention category** Parenting programmes – Webster-Stratton parenting programme Parenting intervention – CBT	Parental autonomy-support - free play 4). no difference SMD-0.89 [-1.70, -0.08] (Hughes 2004). Parental autonomy-support - ring toss 5). no difference SMD-0.26 [-1.04, 0.51] (Hughes 2004). Parenting structure - free play 6). no difference SMD 0.0 [-0.77, 0.77] (Hughes 2004) Parenting structure - ring toss 7). no difference SMD-0.34 [-1.12, 0.44] (Hughes 2004) Parental stress 8). no difference SMD-0.36 [-1.04, 0.31] (Terao 1999). Child behaviour (ECBI, Eyberg Child Behavior Inventory) - intensity score 9). large significant differences favouring the intervention group for intensity of behaviour problems SMD-0.72 [-1.41 to -0.02] and for the number of problems SMD-1.81 [-2.63 to -1.00] (Terao 1999). Child autonomy - free play 10). no difference SMD 0.45 [-0.33, 1.23](Hughes 2004) Child autonomy - ring toss 11). no difference SMD 0.18 [-0.59, 0.95] (Hughes 2004). B. Parenting program vs alternative treatments (CBT or family therapy). child abuse potential:	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Parent-Child Interaction therapy Intervention **Describe intervention** – 1. PCIT programme (1 RCT, Chaffin 2004), aimed to enhance skills and establish daily positive parent-child interaction, followed by command-giving and positive discipline using live coached parent-child dyad sessions. It comprised 6 group-based sessions on increasing parental motivation, followed by clinic-based individual parent-child dyad sessions. Both programmes were delivered over 3 modules (30 sessions). Comparison - Standard community based parenting group. 2. Another PCIT program (1 RCT, Terao 1999), aimed to change patterns of dysfunctional parent-child relationships. The programme was delivered over 14 weekly sessions and comprised behaviour management and communication skills training. Comparison - standard family preservation services. 3. The Webster-Stratton Incredible Years programme (1 RCT, Hughes 2004) - no programme description provided. The programme was delivered over the course of 8 two-hour weekly sessions and was designed to assist parents in learning how to modify their	1). no difference between the two groups SMD 0.03 [-0.42 to 0.48] (Chaffin 2004). Child abuse potential (CAPI) - Rigidity scale 2). no difference SMD 0.41 [-0.04, 0.86] (Chaffin 2004). Child abuse potential (CAPI) - Distress scale 3). no difference SMD -0.11 [-0.56, 0.34] (Chaffin 2004). Child Abuse Potential (CAPI) - Loneliness scale 4). no difference SMD -0.05 [-0.49, 0.40] (Chaffin 2004). Child Abuse Potential (CAPI) Problems with child scale 5). no difference SMD 0.39 [-0.06, 0.85] (Chaffin 2004). Positive parent behaviours (DPICS-II) 6). significant improvement in positive parent behaviour towards the child for the PCIT group SMD 0.50 [0.04, 0.95] (Chaffin 2004). Negative parent behaviours (DPICS-II) 7).significant effect for reduced negative parent behaviour towards the child SMD 0.75 [0.29, 1.22] (Chaffin 2004). Child behaviour (BASC) - externalising 8). no difference SMD 0.06 [-0.39, 0.51] (Chaffin 2004). Child behaviour (BASC) - Internalising 9). no difference SMD -0.02 [-0.47, 0.43] (Chaffin 2004). Parental anger - child report 10). a large significant effect in favour of the CBT group SMD -1.21 [-1.91, -0.51] (Kolko 1996).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	parenting practices following home visits to assess parent-child interaction. Comparison - waiting list (no treatment)control group 4. Cognitive behavioural therapy (CBT) (1 RCT, Kolko 1996) aimed to modify risk factors associated with child physical abuse with an ecologically based family therapy (FT) programme focused on family interaction. Both services comprised 12 one-hour weekly clinic sessions with follow-up home sessions to evaluate progress. Comparison-Family therapy 5-7. Group-based parent-training sessions (3 RCTs, Brunk 1987; Egan 1983,;; Wolfe 1981), aimed to enhance child management skills using a clinic-based multi-systemic family therapy comprising individual family-tailored behavioural management strategies delivered in 8 weekly sessions of 1.5 hours duration. Comparison - clinic-based multi-systemic family therapy (Brunk 1987); or parenting group with stress management training aimed at improving parental emotional control and including relaxation skills training and cognitive restructuring. Comparison - parenting group + stress management, and control (Egan 1983) or parenting group using videotaped vignettes, and self-control using deep	Family problems - child report 11). a large significant effect in favour of CBT SMD -0.96 [-1.64, -0.28] (Kolko 1996). Parental anger - parent report 12). no difference SMD -0.45 [-1.10, 0.19] (Kolko 1996). Family problems - parent report 13). no effects SMD 0.0 [-0.64, 0.64] (Kolko 1996). **Narrative findings - effectiveness** – The authors conclude that there re is insufficient evidence to support the use of parenting programmes to reduce physical abuse or neglect. Only 3 of the included 7 studies assessed the impact of the programme on objective measures of child abuse and one study showed a significant effect, suggesting that parent-child interaction therapy can reduce re-reports of physical abuse. There is limited evidence to show that some parenting programmes may be effective in improving some outcomes that are associated with physically abusive parenting, such as improvement in child behaviours and reduced number of child behaviour problems. There is limited evidence that programmes that provide additional components aimed specifically at addressing factors associated with physically abusive parenting such as anger and stress, are more effective compared with parenting programmes that do not include such components. Significant improvement in positive parent behaviour towards the child was reported in groups receiving PCIT intervention when compared with those in the standard community based parenting group.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	muscle relaxation. Comparison-standard service group. Limited information on details of 'comparison' interventions. Delivered by – No details on who delivered interventions. **Duration, frequency, intensity** –See 'Describe interventions'. **Key components and objectives of intervention** –See 'Describe interventions'. **Location/place of delivery** –Clinic- and home-based. **Describe comparison intervention** –See 'Describe intervention'. Overall, limited information on details of 'comparison' interventions. **Outcomes measured** Incidence of abuse and neglect – Child abuse potential Quality of parenting and parent-child relationships. – Parenting skills, Parenting Behaviours, Parental Competence. Children and young people's health and wellbeing outcomes. – Child Behaviour, Child Autonomy		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Caregiver/parent health and wellbeing outcomes. – Parental stress.		

2. Browne DT, Puente-Duran S, Shlonsky A et al. (2016) A Randomized Trial of Wraparound Facilitation Versus Usual Child Protection Services. Research on Social Work Practice 26: 168-179

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim Study aim: To evaluate whether the addition of a wraparound facilitator to regular child protection services improved child and family functioning over 20 months. **Methodology** RCT. Blocked randomization stratified by site. **Country** Not UK. Canada. **Source of funding** Government - The Ontario Ministry of Children and Youth Services.	**Participants** Caregivers and families. **Sample characteristics** • Age: Parents/family carers: mean age of 32.22 years old; Children mean age: 6.45 years old • Sex: Parents/family carers: mostly mothers. Children: 47.4% were female. • Ethnicity: Not reported. • Religion/belief: Not reported. • Disability: Not reported. • Long term health condition: Not reported. • Sexual orientation: Not reported. • Socioeconomic position: Parents/family carers (mostly mothers): 27.4% ‘‘single’’ parents; 26.4% ‘‘separated,’’ 47.4% had one child living at home at the time of the referral, 27.4% had two, 16.3% had three, and the remaining families had four or more.	**Effect sizes** Effects of Wraparound Versus Usual Care Over 20 Months: A. The effects of intervention (wraparound vs. usual care) and time (baseline and 20-month follow-up) (Table 2) Significant main effect of time for 1. caregiver psychological distress (F=10.88, p=0.001), 2. family resources (F=25.83, p <.001), 3. child impairments (F=31.10, p<.001), indicating that families and children enrolled in the study tended to improve in these areas irrespective of the treatment group. No significant main effect of time for (Table 2) 1. parental stress (F=1.47, p=0.227), 2. developmental milestones (ages and stages) (F=2.55, p=0.116), suggesting neither improvement nor deterioration between intake and follow-up. 3. child strengths was marginally significant (F=3.48, p=0.066). B. There were no significant Intervention x Time interactions, indicating that participants in the experimental and control conditions had improved similarly at the 20-month follow-up, with small to moderate effect sizes (Table 3). Both groups improved in 1. child impairments, d=-0.60 [-0.81 to -.39], 2. caregiver psychological distress, d=-0.33 [-0.52 to -0.13], 3. family resources, d=0.44 [0.27 to 0.62], Post-test treatment effects (Table 3) child impairments, d=0.14 [-0.12 to	**Overall assessment of internal validity** ++ **Overall assessment of external validity** + **Overall validity score** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Type of abuse: Child abuse and neglect. Participants; physical/sexual harm by commission (20%), harm by omission (21.5%), emotional harm/exposure to conflict (32.6%), abandonment/separation (12.6%), caregiver capacity (73.3%), and request for assistance (1.5%). • Looked after or adopted status: N.A. • Unaccompanied asylum seeking, refugee or trafficked children: N.A. **Sample size** • Comparison numbers - Control: CPS only (n=68) • Intervention numbers - Wraparound +CPS (n=68) • Sample size - Total n=135 **Intervention** Describe intervention Wraparound intervention: 'a comprehensive model of care coordination based upon a bioecological model of human development' (p2, citing Bronfenner and Morris, 2006; Bruns, Weather, Suters et al. 2014). Families in the intervention received usual CPS plus assigned a wraparound facilitator. Wraparound facilitators were master's-level social workers who received extensive training and ongoing monitoring from a certified wraparound trainer	0.52] ns Maternal depression d=0.25 [-0.07 to 0.57] ns Family resources, d=-26 [-0.26 to 0.03] ns Parental stress, d=0.10 [-0.19 to 0.40] ns Child strengths, d=-0.24 [-0.37 to 0.29] ns. This suggests no measurable benefit was associated with the intervention. A program fidelity assessment was conducted after the intervention and found ''average'' implementation fidelity (i.e. minimum standards for wraparound) in two of the programme's major components.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	prior to commencement of the trial. The model is a person- and family-driven planning process that is team based. It highlights the importance of flexible adaptation to the unique needs of communities, cultures, care providers, and system partners (p6). 'Before intervention, wraparound facilitators held an initial meeting with children and families in order to identify their hopes and aspirations, assess needs and strengths in multiple domains, and gain an understanding of any individualized or cultural factors that may be important to address during care, especially the social determinants of health such as socioeconomic, housing, and neighbourhood challenges. A personalized wraparound team was created in a chosen support network including friends, extended family members, and both formal and informal supports. The team met regularly, beginning with contact in the 3–5 hr/week range over the first few months, then eventually to 2–3 hr/week once engagement and progress was established. At each meeting, the family's pressing needs were identified and solutions were brainstormed, leading to a concrete written action plan that specifies explicit tasks and roles. Outcomes were tracked		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	throughout the process. Both successes and failures were integrated into future strategies in an iterative and reflective fashion. Throughout the intervention, the three wraparound facilitators met weekly with the expert trainer in order to review cases, address challenges and concerns, and monitor fidelity to wraparound principles' (p6). Integral intervention components identified 10 essential elements for care in wraparound model. These are as follows: 1. promoting family voice and choice in the service plan, 2. providing care that is embedded in the child and family team, 3. drawing upon natural supports in the family's context, 4. collaboration among all formal and informal team members, 5. provision of community-based care, 6. ensuring cultural competence and sensitivity, 7. presence of an individualized care plan, 8. use of a strengths-based model, 9. persistence and problem solving throughout challenges, and 10. outcome-based methods of evaluation (p2). Family voice and choice (freedom to choose and design care plan), team-based (family selected people on team), natural supports (increased support from friends and family), collaboration (family made plan, team members had role in implementation and held one another accountable),		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	community-based (child got involved with community activities, developed access to services and supports), culturally competent (family had time to talk about strengths, and team used understandable language and respected client), individualized (team included people who were not just professionals, resources were available for support and transitions), strength-based (strengths were discussed and supports were connected to child and family abilities), persistence (team came up with new ideas if others weren't working), and outcome-based (went through a process identifying what leads to success) (p6). At the end of treatment, clients completed the WFI (wraparound fidelity index), which assessed these areas in reference to the care they received. **Delivered by** Facilitators who were master's-level social workers who received extensive training. **Delivered to** Families (parents and carers). **Duration, frequency, intensity, etc.** Contact between wraparound facilitators and families 3–5 hr per week		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	range over the first few months, then eventually to 2–3 hr per week once engagement and progress was established. **Key components and objectives of intervention** Wraparound facilitation Integral intervention components identified as family voice and choice (freedom to choose and design care plan), team-based (family selected people on team), natural supports (increased support from friends and family), collaboration (family made plan, team members had role in implementation and held one another accountable), community-based (child got involved with community activities, developed access to services and supports), culturally competent (family had time to talk about strengths, and team used understandable language and respected client), individualized (team included people who were not just professionals, resources were available for support and transitions), strength-based (strengths were discussed and supports were connected to child and family abilities), persistence (team came up with new ideas if others weren't working), and outcome-based		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	(went through a process identifying what leads to success) (p6). **Content/session titles** Wrap-around Facilitation. **Location/place of delivery** Not clear, likely to be at home. **Describe comparison intervention** Control group received CPS care as usual: delivered by usual care workers had a range of certifications and backgrounds. Current standards indicate that workers and families must be in direct contact at least monthly, with formal reassessment occurring every 6 months. Over the course of the current 20-month trial, this translated into 20 meetings and 3 formal revaluations. The control group also filled out the WFI in response to the care they received (p6). **Outcomes measured** Children and young people's health and wellbeing outcomes. Functional impairments, Behavioral and emotional strengths, Developmental milestones. Caregiver/parent health and wellbeing outcomes. Caregiver psychological distress, adequacy of family resources		

3. DePrince AP, Chu AT, Labus J S et al. (2015) Testing Two Approaches to Revictimization Prevention Among Adolescent Girls in the Child Welfare System. Journal of Adolescent Health 56: S33-S39

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim Study aim: To 'compare two interventions designed to decrease revictimisation in a diverse sample of adolescent child-welfare involved girls' (pS33). **Methodology** RCT including cluster. **Country** Not UK. USA - Colorado **Source of funding** Government Other - University of Denver	**Participants** Children and young people - Adolescent girls (aged 12-19) with histories of child neglect or abuse. **Sample characteristics** • Age - Range 12 to 19, mean 15.85 (sd=1.58). • Sex - All female. • Ethnicity - White/Caucasian 36%, Black/African American 36%, American Indian/Native Alaskan/Native American 7%, Asian/Asian-American 3%, Other 18%. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - 77% reported as heterosexual/straight. • Socioeconomic position - Not reported. • Type of abuse - Witnessing domestic violence 69%, neglect 43% sexual abuse 40%, physical abuse 37%, emotional psychological abuse 35%. Mean age of onset 5.56 years (sd=4.39), average number of perpetrators was 2.51 (sd=2.00). 63% reported victimisation by peer perpetrators in addition	**Effect sizes** Incidence of abuse and neglect. Note: Only contrasts between the two intervention groups are reported, not contrasts with 'no treatment group'. 'No treatment group' not randomly allocated. 1) Sexual revictimisation No significant differences between the odds ratios for sexual revictimisation were observed between the social learning/feminist theory intervention and executive function/risk detection intervention. 2) Physical revictimisation No significant differences between the odds ratios for physical revictimisation were observed between the social learning/feminist theory intervention and executive function/risk detection intervention.	**Overall assessment of internal validity** + **Overall assessment of external validity** ++ US study, but background services likely to be the same for this type of intervention as in UK. **Overall validity score** + Limitations include no intent to treat analysis, and creation of a 'no treatment' comparison group of those who did not attend any sessions.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	to victimisation by adults. • Looked after or adopted status - Biological/natural family 27%, foster home 23%, group home 17%, residential treatment facility 12%, independent living program 4%, with relatives 6%, on own 3%, with adoptive family 3%, declined to answer 4%. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** Comparison numbers: n=67 (no treatment group have not been included in this data extraction). Intervention numbers - n=67 Sample size: n=134 **Describe intervention** Intervention based on social learning/feminist theory. From this perspective, children exposed to violence may learn that this is acceptable, and fail to learn coping skills. Uses Youth Relationships Manual (Wolfe et al. 1996) targeting 1) power in relationship violence, 2) developing skills to build healthy relationships and recognise abuse 3) developing skills to respond to pressures that can lead to violence		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	4) increasing competency. Intervention comprised 12 weekly intervention group meetings, co-facilitated by graduate students. Each session lasted 1.5 hours. Receipt of newsletters with local and telephone resources for violence. Individual onward referrals made as appropriate. **Describe comparison intervention** Based on a risk detection and executive function perspective. This involves 'noticing and responding to external...and internal...danger cues in intimate relationships' (p. S34). Working on cognitive skills to support this including focusing attention, taking on board new information, thinking flexibly about solutions and planning and initiating actions. As for social learning/feminist intervention, risk detection intervention comprised 12 weekly intervention group meetings, co-facilitated by graduate students. Each session lasted 1.5 hours. Receipt of newsletters with local and telephone resources for violence. Individual onward referrals made as appropriate. **Outcomes measured** Incidence of abuse and neglect 1) Violence exposure (revictimisation) assessed using Traumatic Events		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Screening Inventory (TESI) Child version at baseline, immediately after 12 week intervention, 2 months and 6 months. 2) Dating violence measured using Conflict in Adolescent Dating Relationships Inventory (CADRI).		

2. Donohue B, Azrin NH, Bradshaw K et al. (2014) A controlled evaluation of family behavior therapy in concurrent child neglect and drug abuse. Journal of Consulting and Clinical Psychology 82: 706–20

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim – The study examines 'the effects of Family Behavioral Therapy as compared to treatment as usual community-based services (TAU)' (p708). **Methodology** – Design is a 2 (treatment type: FBT, TAU) x 2 (neglect type: neglect due to foetus/child being exposed to drugs, other child neglect) x 3 (time: baseline, 6 months postrandomisation and 10 months postrandomisation) mixed model experimental design.	**Participants** Caregivers and families. – Mothers (this is specified by the research as opposed to 'parents') who: a) Had been reported to the Department of Family Services for child neglect b) Were living with the child who were subject to the neglect referral c) Were identified as using illicit drugs during the 4 months prior to the referral d) Were displaying symptoms consistent with illicit drug abuse or dependence at the time of referral e) Had at least one other adult individual willing to participate in their treatment f) Whose primary reason for referral was not due to sexual abuse perpetration or domestic violence. The authors also distinguish between parents for whom neglect consists of exposing their child to drug use in utero or in childhood,	**Effect sizes** Risk of abuse and neglect (Note: We do not report main effects of time, or time x neglect type interactions here as they are not relevant to the review question regarding effectiveness of the intervention.) – There was a significant time x treatment x neglect type interaction from baseline to 6 months postrandomisation - F(1, 68)=5.977, p=0.009, partial eta-squared=0.081 - and baseline to 10 months postrandomisation - F(1, 68)=3.329, p=0.04, partial eta squared =0.047. 'Post hoc analyses indicated that FBT mothers of drug exposed children reduced their maltreatment potential more than FBT mothers of drug-exposed children and TAU mothers (p<0.05)' (p71). (Note: the paper does not give ANOVA result for treatment x time interaction, so the overall effectiveness of FBT compared to TAU across both categories of neglect is not clear). The overall effect sizes for CAPI are described as 'medium' for FBT and 'small' for TAU (FBT 6 months 0.41 [-.11,.94], 10 months 0.41 [-.10, .92]; TAU 6 months	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall validity score:** - A key methodological flaw is the failure to report cell sizes for drug-exposed versus non-drug exposed families. This makes it difficult to judge the validity of the statistical analysis. There is also lack of clarity regarding data imputation methods for intent to treat analysis

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Country Not UK – US. **Source of funding** Other. – National Institutes of Health (USA).	and those who neglect 'for other reasons' (p708). Later in the paper, these reasons are stated to include lack of supervision, emotional, medical, environmental or physical neglect (p.716). It's unclear how many parents were in each category. **Sample characteristics** Age – Mothers: Overall mean age 29.04 (sd=8.07); FBT mean age 29.63 (7.65); TAU mean age 28.49 (8.51) Children: Overall mean age 3.92 (3.73); FBT mean age 4.20 (4.06); TAU mean age 3.65 (3.42). Sex – All caregivers appear to be mothers. Sex of children not provided. Ethnicity – Caucasian – Overall 34 (47.2%); FBT 14 (40.0%); TAU 20 (54.1%) Black/African American - Overall 18 (25.0%); FBT 10 (28.6%); TAU 8 (21.6%) Hispanic/Latino - Overall 8 (11.1%); FBT 6 (17.1%); TAU 2 (5.4%) Asian American - Overall 2 (2.8%); FBT 2 (5.7%); TAU 0 (0.0%) Pacific Islander - Overall 2 (2.8%); FBT 1 (2.9%); TAU 1 (2.7%) Other - Overall 5 (6.9%); FBT 0 (0.0%); TAU 5 (13.5%)	0.23[-.25, .70], 10 months 0.27 [-0.21, 0.74). (NB effect sizes appear to measure the size of the effect comparing mean scores at baseline with 6 and 10 month measures, rather than comparing the mean scores between the 2 interventions). **Children and young people's health and wellbeing outcomes** – Time spent by child in Department of Family Services custody There was no difference between FBT and TAU in terms of days spent by children in DFS custody. In both conditions, children were significantly more likely to spend time in DFS custody at 6 months postrandomisation (F(1, 68)=7.625, p=0.004, partial eta squared=1.01), but not at 10 months postrandomisation. This is reflected in the effect sizes for both interventions (FBT 6 months -0.24 [-.76, .28], FBT 10 months -0.04 [-.55, .47]; TAU 6 months -0.28 [-.76, .19], TAU 10 months -.12 [-.59, .36]. **Caregiver/parent health and wellbeing outcomes** – Marijuana use: There were no significant differences in parental marijuana use between conditions (p values > 0.05). Effect sizes for both FBT and TAU were medium. (NB effect sizes appear to measure the size of the effect comparing mean scores at baseline with 6 and 10 month measures, rather than comparing the mean scores between the two interventions). (FBT 6 months 0.74 [.22-1.27], FBT 10 months 0.63 [.12, 1.15]; TAU 6 months 0.55 [.07, 1.04], TAU 10 months 0.53 [.06, 1.01]).	(or indeed if imputation was used).

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Religion/belief – Not reported. Disability – Not reported. Long term health condition – Not reported. Sexual orientation – Not reported. Socioeconomic position – Employment Unemployed - Overall 63 (87.5%); FBT 28 (80.0%); TAU 35 (94.6%) Employed full time - Overall 5 (6.9%); FBT 3 (8.6%); TAU 2 (5.4%) Employed part time - Overall 4 (5.6%); FBT 4 (11.4%); TAU 0 (0.0%) Education Less than high school - Overall 36 (50.0%); FBT 19 (54.3%); TAU 17 (45.9%) High school/equivalent - Overall 32 (44.4%); FBT 15 (42.9%); TAU 17 (45.9%) Degree - Overall 4 (5.6%); FBT 1 (2.9%); TAU 3 (8.1%). Type of abuse – Included families were those who had had referrals to CPS on grounds of neglect. Looked after or adopted status – Not reported. Unaccompanied asylum seeking, refugee or trafficked children – Not reported.	'Hard drug' use (illicit drugs other than marijuana) There was a significant time x treatment x neglect interaction from baseline to 6 months - F(1, 68)=5.577, p=0.015, partial eta squared=0.076 - and from baseline to 10 months - F(1, 68)=8.148, p=0.003, partial eta squared =0.107. 'Post hoc analysis indicated that mothers of non-drug exposed children in FBT demonstrated significant decreases in hard drug use, as compared with mothers of drug-exposed children in FBT and mothers of non-drug exposed children in TAU (p<0.05). Mothers of drug-exposed children in TAU demonstrated significant decreases in hard drug use, as compared with mothers of drug-exposed children in FBT and mothers of non-drug-exposed children in TAU' (pp713-714). Effect sizes for both FBT and TAU were small to medium (FBT 6 months 0.41 [-.11, 0.93], FBT 10 months 0.39 [-.12, 0.90]; TAU 0.31 [-.16, .79], TAU 10 months 0.45 [-.2, 0.93]. Risk of HIV transmission: There was a significant treatment x time interaction for risk of HIV transmission - F(1, 68) = 4.014, p=0.03, partial eta squared = 0.056, suggesting that Participants in FBT showed greater improvements in risk of HIV transmission than the comparison group. The hypothesised time x neglect type x treatment interactions were not observed (p values >0.05). Effect sizes for risk of HIV transmission for FBT were medium, whereas those in TAU showed almost no effect from baseline- to post, and only a small effect from baseline to follow up (FBT post 0.33 [-0.19, 0.85], FBT follow-up 0.33 [-0.18, 0.84]; TAU post 0.00 [-0.48, 0.48], TAU follow-up 0.24 [-.24, 0.71]). (Reviewer note: assume that post = 6 months and follow-	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Sample size Comparison numbers – n=37 Intervention number – n=35 **Sample size** – n=72 **Intervention category** Family Behaviour Therapy. **Intervention** **Describe intervention** – Intervention has been adapted from family behaviour therapy, a 'comprehensive outpatient treatment equipped to manage substance disorders' (p.709). The authors note that FBT 'emphasises cognitive and behavioral skill development through behavioral role-playing, therapeutic assignments and utilisation of family support systems' (p709). **Delivered by** – Eleven providers with no previous experience of implementing FBT. Qualifications ranged from bachelor's level to doctorate. Providers received approximately 16 hours of training and attended 90 to 120 minutes of weekly group supervision throughout the study.	up = 10 months, although this is not clear in the paper). Hours of employment: There was a significant time x treatment interaction in number of hours employed from baseline to 6 months - F(1, 68)=3.868, p=0.027, partial eta squared=0.054 - and from baseline to 10 months - F(1, 68) = 3.549, p=0.032, partial eta squared = 0.05, in favour of FBT. The effect sizes for this variable for FBT are small, and for TAU the effect size is small at 6 months and negligible at 10 months (FBT 6 months -.18 [-0.70, 0.34], 10 months -0.30 [-0.80, 0.21]; TAU 6 months 0.23 [-0.25, 0.71], 10 months 0.04 [-.43, 0.52]. Alcohol intoxication: There was no significant difference between levels of alcohol intoxication between FBT and TAU (no time x treatment interaction, and no time x treatment x neglect type interaction) (p values >0.05). There were small to medium effect sizes for both conditions (FBT 6 months 0.31 [-.21, 0.83], FBT 10 months 0.37 [-.14, .88]; TAU 6 months 0.11 [-.37, .59], TAU 10 months .33 [-.14, 0.81]. Incarceration: There was a marginally significant time x treatment interaction from baseline to 6 months postrandomisation - F(1, 68) = 2.554, p = 0.058, partial eta squared = 0.036, with FBT Participants spending less time incarcerated than TAU Participants. The time x treatment x neglect type interaction was not significant. There was no effect of FBT on incarceration. The effect size for TAU was medium, but the direction of the effect is in terms of increased number of days incarceration. (FBT 6 months 0.02 [-.50, .54], FBT 10	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Delivered to – Mothers who had been reported to the Department of Family Services for child neglect and were involved in illicit drug use. **Duration, frequency, intensity** – Mothers attended an average of 14.9 meetings. **Key components and objectives of intervention** – Intervention has been adapted from family behaviour therapy, a 'comprehensive outpatient treatment equipped to manage substance disorders' (p.709). The authors note that FBT 'emphasises cognitive and behavioral skill development through behavioral role-playing, therapeutic assignments and utilisation of family support systems' (p.709). It involves implementing the following components a) helping significant others to provide family-derived rewards for prosocial target behaviours; n) communication skills; c) stimulus control interventions to promote spending time with individuals and situations not involved substance misuse and other problem behaviors; d) self-control methods to manage drug cravings; e) skills training specific to employment. In this study, FBT was	months 0.03 [-.48, .54]; TAU 6 months -0.40 [-.88, .08], TAU 10 months .0.35 [-.83, 0.12]). **Narrative findings** Narrative findings – effectiveness **– Risk of abuse and neglect (measured using Child Abuse Potential Inventory)** The authors conclude that 'FBT was more effective than TAU in reducing child maltreatment potential in mothers of non-drug-exposed children from baseline to 6 and 10 months postrandomization' (p715). 'Non-drug-exposed' refers to forms of neglect which do not involve exposure to drugs in utero or in childhood (for example, lack of supervision). However, FBT was not more effective than TAU for mothers of drug-exposed children. The authors suggest this may be due to older age and greater family stability of non-drug-exposing parents, meaning that they were better able to engage in treatment. **Children and young people's health and wellbeing outcomes (measured by days in DFS custody).** There was no difference between FBT and TAU in terms of days spent by children in DFS custody. In both conditions, children were significantly more likely to spend time in DFS custody at 6 months postrandomisation compared to baseline. **Caregiver/parent health and wellbeing outcomes** Marijuana use: There were no significant differences in parental marijuana use between FBT and TAU. Effect sizes for this variable both FBT and TAU were medium.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	adapted for use with families by seeing service users at home; treatment session increased from 60 to 75 minutes; duration of treatment increased from 4 to 6 months; target number of treatment sessions increased from 15 to 20; the following intervention components added: a) identification of home hazards, b) improving financial management skills c) teaching mothers how to reinforce good behaviours in their children, d) teaching mothers to react to emergent conditions affecting their families, e) HIV and STD prevention. **Content/session titles** – See above. **Location/place of delivery** – Participants' homes. **Describe comparison intervention** – Comparison was 'treatment as usual' (TAU) which comprised 'a variety of services that vary according to provider qualifications, duration, intensity and type of services offered' (p711). Caseworkers referred families to different services, depending on need, problem severity, motivation and the availability of services. Referrals included child placement, crisis intervention services, family services (e.g. family therapy), caregiver services (e.g.	'Hard drug' use (illicit drugs other than marijuana) - Mothers of non-drug exposed children in FBT demonstrated significant decreases in hard drug use, as compared with mothers of drug-exposed children in FBT and mothers of non-drug exposed children in TAU. However, mothers of drug-exposed children in TAU demonstrated significant decreases in hard drug use, as compared with mothers of drug-exposed children in FBT and mothers of non-drug-exposed children in TAU. Within-subjects effect sizes for this variable for both FBT and TAU were small to medium. Risk of HIV transmission - Participants in FBT showed greater improvements in risk of HIV transmission than the comparison group. However, the authors describe these improvements as 'shortlived' (p.716). Within-subjects effect sizes for risk of HIV transmission for FBT were medium, whereas those in TAU showed almost no effect from baseline- to post, and only a small effect from baseline to follow up. (Reviewer note: assume that post = 6 months and follow-up = 10 months, although this is not clear in the paper). Hours of employment - Participants in the FBT group showed a greater increase in hours of employment compared to TAU, with a small between-subjects effect size. The within-subjects effect sizes for this variable for FBT are small, and for TAU the within-subjects effect size is small at 6 months and negligible at 10 months. Alcohol intoxication - There was no significant difference between levels of alcohol intoxication between	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	counselling), child services, and other 'miscellaneous' services. **Outcomes measured** Risk of abuse and neglect – Potential for maltreatment was measured using the Child Abuse Potential Inventory (CAPI; Milner 1986). This measure has recognised cross-cultural validity, internal consistency for subscale and total scores and differential validity (Walker and Davies 2010). Children and young people's health and wellbeing outcomes – Participant and significant other Timeline Follow-Back was used to assess the number of days Participants' children had been in DFS custody (note: unsure whether this would reflect abuse taking place, or whether children could be placed in care of Department of Family Services for other reasons). Caregiver/parent health and wellbeing outcomes – Drug use (frequency of days of use of marijuana and other illicit drugs other than marijuana - hard drugs) measured during the four months prior to assessment using the Timeline Follow-Back (TLFB; Sobell and Sobell	FBT and TAU. There were small to medium within-subjects effect sizes for both conditions. Incarceration - FBT Participants spent less time incarcerated than TAU Participants. This difference was marginally significant. There was no within-subjects effect of FBT on incarceration, and a medium within-subjects effect on incarceration for TAU, in the direction of increased number of days incarceration. The authors interpret this as indicating that FBT may have helped to prevent future incarceration.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	1992). The TLFB is completed both by the parent and their 'primary adult significant other' (p711). Drug use also measured using urinalysis testing. Where there were conflicts between the three measures, the measure indicating greatest substance use during the previous four months was used. Risk of HIV transmission was measured using the Total Risk Scale of the Risk Assessment Battery (Metzger et al. 1990). Internal consistency of this scale is poor to good (Cronbach alpha 0.42 to 0.82). Participant and partner TLFB were used to assess: - hours employed - days using alcohol - days incarcerated.		

4. Fantuzzo J, Manz P, Atkins M et al. (2005) Peer-mediated treatment of socially withdrawn maltreated preschool children: Cultivating natural community resources. Journal of Clinical Child and Adolescent Psychology 34: 320–25

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim – To evaluate the effectiveness of Resilient Peer Treatment, a '… peer-mediated, classroom-based intervention for socially withdrawn, maltreated preschool children' (p320). The intervention had previously been found	**Participants** **Children and young people** – Recipients of the interventions were socially withdrawn African American children from 40 participating Head Start classrooms in the northeast of the USA. Eligibility was determined on the basis of teacher ratings of prosocial peer interactions and independent verification of these. Maltreatment was substantiated for around half of this	**Effect sizes** **Children and young people's health and wellbeing outcomes** Dyadic play interactions in play corners (observed, two weeks pre and two weeks post intervention) Two-way analysis of covariance (treatment x maltreatment) - collaborative play (at post-test, controlling for pre-test scores using a Bonferroni correction): There was a main effect of treatment group, in favour of the intervention, with large effect size: $F(1, 77)=39.1$, $p<.0001$, $\eta 2 = .36$. Maltreatment status: No significant	**Overall assessment of internal validity** - **Overall assessment of external validity** ++ **Overall validity score** - – The use of coding systems and scales with unclear reliability and

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
to be effective in a treatment setting and the authors aimed to determine whether this could be transferred to the classroom setting. **Methodology** RCT **Country** Not UK. – USA - northeast. **Source of funding** Not reported.	group. A second group of children identified as having the highest levels of prosocial peer play interactions were chosen to be 'play buddies'. **Caregivers and families** – Parent volunteers identified by teachers and 'parent involvement staff members' were invited to take on the role of 'play supporters'. **Professionals/practitioners** – Teachers working in one of 40 participating Head Start classrooms in the northeast of the USA. **Sample characteristics** • Age -The mean age of children who received the interventions was 4.35 years (SD= .47). Age of play buddies and play supporters is not reported. • Sex - 50% of children who received the interventions were male. Gender of play buddies and play supporters is not reported. • Ethnicity - All children who received the interventions were African American. Ethnicity of play buddies and play supporters is not reported. • Religion/belief– Not reported. • Disability– Not reported. • Long term health condition– Not reported. • Sexual orientation– Not reported.	effect (statistical data not provided). Treatment x maltreatment interaction: No significant effect (statistical data not provided). Two-way analysis of covariance (treatment x maltreatment) – solitary play (at post-test, controlling for pre-test scores using a Bonferroni correction): There was a main effect of treatment group, in favour of the intervention, with medium to large effect size: F(1, 77)=13.7, p<.0001, η2 = .15. Maltreatment status: No significant effect (statistical data not provided). Treatment x maltreatment interaction: No significant effect (statistical data not provided). Two-way analysis of covariance (treatment x maltreatment) – social attention (at post-test, controlling for pre-test scores using a Bonferroni correction): Treatment group: No significant effect (statistical data not provided). Maltreatment status: No significant effect (statistical data not provided). Treatment x maltreatment interaction: No significant effect (statistical data not provided). Two-way analysis of covariance (treatment x maltreatment) – associative play (at post-test, controlling for pre-test scores using a Bonferroni correction): Treatment group: No significant effect (statistical data not provided). Maltreatment status: No significant effect (statistical data not provided). Treatment x maltreatment interaction: No significant effect (statistical data not provided). Free-play observations at two weeks post-intervention – Collaborative play (observed during free-play at two weeks post-intervention) - Two-way analysis of variance (using Bonferroni correction) – There was a main effect of treatment group, in favour of the intervention, with medium to large effect size: F(1, 78) = 19.0, p <	validity and a very short follow-up (2 weeks) mean that it is difficult to be confident in the findings of this study.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Socioeconomic position– 72% of children who received the interventions lived in single-female households; 32%of the parents of these children had no high-school education; 74% cent of these parents were unemployed. Socioeconomic data for play buddies and play supporters is not reported. • Type of abuse– Maltreatment was substantiated for 37 of the 82 children who received the interventions. Twelve children had been physically abused, 18 had been physically neglected and seven had been both physically abused and physically neglected. The authors report that the first documented incident, occurred, on average between the ages of two and three and there were on average 1.4 incidents per child. This ranged from minor to moderate injuries. • Looked after or adopted status– Not reported. • Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size** Comparison numbers – n=44. Maltreated n=19, non-maltreated n=25. Intervention number	.0001, η2=.19. Maltreatment status: No significant effect (statistical data not provided). Treatment x maltreatment interaction: No significant effect (statistical data not provided). Associative play (observed during free-play at 2 weeks post-intervention) – 2-way analysis of variance (using Bonferroni correction) - Treatment group: No significant effect (statistical data not provided). Maltreatment status: No significant effect (statistical data not provided). Treatment x maltreatment interaction: No significant effect (statistical data not provided). Social attention (observed during free-play at two weeks post-intervention) – 2-way analysis of variance (using Bonferroni correction) – Treatment group: No significant effect (statistical data not provided). Maltreatment status: No significant effect (statistical data not provided). Treatment x maltreatment interaction: No significant effect (statistical data not provided). Solitary play (observed during free-play at two weeks post-intervention) – 2-way analysis of variance (using Bonferroni correction) - There was a main effect of treatment group, in favour of the intervention, with medium to large effect size: F(1, 78)=12.4, p<.001, η2=.14. Maltreatment status: No significant effect (statistical data not provided). Treatment x maltreatment interaction: No significant effect (statistical data not provided). Penn Interactive Peer Play Scale (Teacher completed at two weeks post-intervention) multivariate analysis (NB raw data not reported): Significant effect for treatment group: Wilks' Λ=.79, F(3, 76)=6.9, p<.001. Maltreatment status: No significant effect (statistical data not provided). Treatment x maltreatment interaction: No significant effect (statistical data not provided).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	– n=38. Maltreated n=18, non-maltreated n=20. **Sample size** – n=82. **Intervention category** Other – Resilient Peer Treatment **Intervention** **Describe intervention** – Resilient Peer Treatment is a child focused '… peer-mediated, classroom-based intervention for socially withdrawn, maltreated pre-school children (p320). **Delivered by** – Participating children interact in play sessions with play buddies (classmates of the participating child who are assessed as showing high levels of prosocial peer play). The session is facilitated by a play supporter (parent volunteer identified by teachers and parents involved in the design of the study as being supportive and nurturing). **Delivered to** – Socially withdrawn African American children from participating Head Start classes in the northeast of the USA. **Duration, frequency, intensity**	Penn Interactive Peer Play Scale (Teacher completed at two weeks post-intervention) univariate analysis (treatment group): Play interaction subscale: There was a main effect of treatment group, in favour of the intervention, with a medium to large effect size, $F(1, 78)=15.4$, $p<.001$, $\eta 2=.16$. Play disruption subscale: There was a main effect of treatment group, in favour of the intervention, with a medium to large effect size, $F(1, 78)=6.0$, $p<.05$, $\eta 2=.07$. Play disconnection subscale: There was a main effect of treatment group, in favour of the intervention, with a medium to large effect size, $F(1, 78)=12.2$, $p<.001$, $\eta 2=.14$. Social Skills Rating System (Teacher completed at two weeks post-intervention) – Social skills scale - multivariate analysis: Significant effect for treatment group: Wilks' $\Lambda=.77$, $F(3, 76)=7.4$, $p<.001$. Maltreatment status: No significant effect (statistical data not provided). Treatment x maltreatment interaction: No significant effect (statistical data not provided). Social Skills Rating System (Teacher completed at two weeks post-intervention) – Social skills scale - univariate analysis: Self-control subscale: There was a main effect of treatment group, in favour of the intervention, with a medium to large effect size, $F(1, 78)=13.4$, $p<.05$, $\eta 2=.15$. Interpersonal skills subscale: There was a main effect of treatment group, in favour of the intervention, with a medium to large effect size, $F(1, 78) = 18.18.77$ [sic], $p<.001$, $\eta 2=.19$. NB There appears to be an error in reporting the F value in relation to this measure. Verbal assertion subscale: No significant differences between groups (statistical data not provided). Social Skills Rating System (Teacher completed at two weeks post-intervention) – Problem behaviors scale - multivariate analysis: Significant effect for treatment	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	– 15 sessions over a two month period (three sessions planned per week). **Key components and objectives of intervention** – The main aim of the intervention is to improve the social competence of withdrawn, maltreated pre-school children by providing opportunities for regular positive play interactions with peers (play buddies) displaying high levels of social functioning. The intervention consists of a number of play sessions with a peer which are supported by an adult volunteer (play supporter). Play supporters facilitate sessions by setting up the play corner in the classroom e.g. putting out toys (usually available in Head Start classrooms); preparing the play buddy for the session (i.e. discussing specific activities which led to positive interactions; observing the play session and providing supportive comments to participating child and their play buddy regarding their interactions. **Content/session titles** – N/A. **Location/place of delivery** – Classroom – play corner. **Describe comparison intervention**	group: Wilks' Λ=.80, F(2, 77)=8.7, p<.001. Maltreatment status: No significant effect (statistical data not provided). Treatment x maltreatment interaction: No significant effect (statistical data not provided). Social Skills Rating System (Teacher completed at two weeks post-intervention) – Problem behaviors scale - univariate analysis: Internalising: There was a main effect of treatment group, in favour of the intervention, with a medium to large effect size, F(1, 78)=12.7, p<.001, η2=.14 Externalizing: There was a main effect of treatment group, in favour of the intervention, with a small to medium effect size, F(1, 39)=8.8, p<.001, η2=.10 Chi-square analysis of group differences on internalizing and externalizing subscales of the Social Skills Rating System Problem behaviors scale – ('… to determine if there were group differences in the number of Participants whose scores were at least 1 standard deviation above the mean range on these behaviour problem scales (T 60)' (p324) Internalizing: Significantly more children in the control group had scores in the higher range on this scale than those in the intervention group, chi-square (1)=7.9, p<.01. Externalizing: Significantly more children in the control group had scores in the higher range on this scale than those in the intervention group, chi-square (1)=5.0, p<.05. **Narrative findings** Narrative findings - effectiveness **Children and young people's health and wellbeing outcomes** Dyadic play interactions in play corners (observed, two weeks pre and two weeks post intervention)	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	– Attention control. Delivered by - Participating children spend time with a peer (assessed as showing average levels of interactive play ability). These interactions are supervised by play supporters (parent volunteer identified by teachers and parents involved in the design of the study as being supportive and nurturing). The authors note that 'Play conditions were identical to the children in the RPT condition, except that they were not paired with a Play Buddy and the Play Supporter … only supervised the play (no prompts or encouragement of play' (p323). Delivered to - Socially withdrawn African American children from participating Head Start classes in the northeast of the USA. **Duration, frequency, intensity** - 15 sessions over a 2 month period (three sessions planned per week). **Key components and objectives of intervention** – The authors report that this '… condition was designed to control for the extra attention of being paired with a peer and spending time with this peer in a play corner under the supervision of a parent volunteer…' (p323). Children in this group used the same 'play corner' and toys	- Collaborative play: There was a significant effect for treatment group, with children randomised to the experimental condition showing higher levels of collaborative play than those in the control group. There was no significant effect for maltreatment status or treatment x maltreatment status interaction on levels of collaborative play. Solitary play: There was a significant effect for treatment group, with children randomised to the experimental condition showing lower levels of solitary play than those in the control group, this showed a small to medium effect size. There was no significant effect for maltreatment status or treatment x maltreatment status interaction on levels of solitary play, meaning that the intervention was not significantly more effective for maltreated compared to non-maltreated children. Social attention: There were no significant effects for treatment group, maltreatment status or treatment x maltreatment status interaction on levels of social attention. Associative play: There were no significant effects for treatment group, maltreatment status or treatment x maltreatment status interaction on levels of associative play. Free-play observations at two weeks post-intervention – Collaborative play: There was a significant effect for treatment group, with children randomised to the experimental condition showing higher levels of collaborative play than those in the control group, this showed a medium to large effect size. There was no significant effect for maltreatment status or treatment x maltreatment status interaction on levels of collaborative play, meaning that the intervention was not significantly more effective for maltreated compared to non-maltreated children.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	as those used in the experimental condition but were not paired with a play buddy. Play supporters assigned to this group provided supervision only and did not offer prompts or encouragement. **Content/session titles** – N/A. **Location/place of delivery** – Classroom – play corner. **Outcomes measured** Children and young people's health and wellbeing outcomes – Social interaction with peers was measured using the Interactive Peer Play observational Coding System (Fantuzzo et al. 1996). Videotaped interactions were coded by researchers into 1 of 4 categories (collaborative play, associative play, social attention). Classroom play with peers was measured using the Penn Interactive Peer Play Scale (Fantuzzo et al. 1998; Fantuzzo et al. 1995) (teacher completed) and analysed using two-way multivariate analyses and univariate analyses (play interaction, play disruption and play disconnection subscales). Children's social behaviour in the classroom was measured using the teacher completed Social Skills Rating System (Gresham and Elliot	Associative play: There were no significant effects for treatment group, maltreatment status or treatment x maltreatment status interaction on levels of associative play. Social attention: There were no significant effects for treatment group, maltreatment status or treatment x maltreatment status interaction on levels of social attention. Solitary play (observed during free-play at two weeks post-intervention) - There was a significant effect for treatment group, with children randomised to the experimental condition showing lower levels of solitary play than those in the control group, this showed a medium to large effect size. There was no significant effect for maltreatment status or treatment x maltreatment status interaction on levels of solitary play, meaning that the intervention was not significantly more effective for maltreated compared to non-maltreated children. Penn Interactive Peer Play Scale (Teacher completed at two weeks post-intervention) multivariate analysis: There was a significant effect for treatment group on the Penn Interactive Peer Play Scale but it is not clear whether this was in favour of the experimental or control condition. No significant effects were found for maltreatment status or treatment x maltreatment status interaction. Penn Interactive Peer Play Scale (Teacher completed at two weeks post-intervention) univariate analysis (treatment group): Significant effects were found for treatment group, with children randomised to the experimental condition being rated significantly higher on the play interaction subscale; and significantly lower on the play disruption (small to medium effect size) and play disconnection subscales (medium	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	1990) and analysed using two-way multivariate analyses and univariate analyses (Social skills scale - self-control, interpersonal skills, verbal assertion subscales; problem behaviors scale – internalising and externalizing subscales).	to large subscales) of the Penn Interactive Peer Play Scale. Social Skills Rating System (Teacher completed at two weeks post-intervention) – Social skills scale - multivariate analysis: There was a significant effect for treatment group on the Social Skills scale of the Social Skills Rating System but it is not clear whether this was in favour of the experimental or control condition. No significant effects were found for maltreatment status or treatment x maltreatment status interaction. Social Skills Rating System (Teacher completed at two weeks post-intervention) – Social skills scale - univariate analysis: Significant effects were found for treatment group, with children randomised to the experimental condition being rated significantly higher on the self-control and interpersonal skills subscale (medium to large effect sizes) of the Social Skills Rating System. No significant effects for treatment group were found on the verbal assertion subscale. Social Skills Rating System (Teacher completed at two weeks post-intervention) Problem behaviors scale - multivariate analysis: There was a significant effect for treatment group on the Problem Behaviors scale of the Social Skills Rating System but it is not clear whether this was in favour of the experimental or control condition. No significant effects were found for maltreatment status or treatment x maltreatment status interaction. Social Skills Rating System (Teacher completed at two weeks post-intervention) Problem behaviors scale - univariate analysis: Significant effects were found for treatment group, with children randomised to the experimental condition being	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		rated as displaying significantly lower levels of internalizing (medium to large effect size) and externalizing behaviours (small to medium effect size) on the Problem Behaviors scale of the Social Skills Rating System. Chi square analyses showed that significantly greater numbers of children in the control condition had scores in the higher ranges on both of these measures.	

5. Fisher PA, Nurraston B, Pears KC (2005) The Early Intervention Foster Care Program: permanent placement outcomes from a randomized trial. Child Maltreatment 10: 61–71

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: This paper forms part of a larger RCT exploring effectiveness of Early Intervention Foster Care Program. This paper reports permanent placement outcomes. **Methodology:** RCT including cluster. Analysis of data from 54 children who were placed in a permanent placement during the EIFC study (out of a total of 90). **Country:** Not UK. US.	**Participants** **Children and young people** Total 54 children and young people who had been fostered, and were placed in a permanent placement during the study. 'Permanent placements were defined as the final nonfoster care placement for the child' (p65). This included reunification with biological parent, relative adoption and non-relative adoption. **Sample characteristics** • Age - Mean age at study start: Early Intervention Foster Care (EIFC)=4.50 (0.86), Regular Foster Care (RFC)=4.22 (0.74).	**Effect sizes** **Service outcomes** 1) Success of permanent placement There was a significant difference in favour of the intervention in rates of failed placement (EIFC 10%; RFC 36%; chi-square=5.11, p=0.02. Effect size calculated by reviewing team r=0.31).). Two children in the RFC group also experienced a second permanent placement breakdown, whereas none in the EIFC group did. 2) Regression model A Cox regression model explaining the variance in failure rates in permanent placement was constructed. The regression model suggests that some of the difference in rates of failed placement across the 2 conditions may have been moderated by other factors, most notably multiple placements during the study.	**Overall assessment of internal validity** + **Overall assessment of external validity** + Unclear to what extent US training for foster carers is similar to UK. **Overall validity score** + Relatively small sample size (n=54). Focuses on placement outcomes only.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Source of funding: Government - US Public Health Service.	• Sex - EIFC: 66% male, RFC: 60% male. • Ethnicity - EIFC: White 79%, Native American 3%, Hispanic or Latino 18% RFC: White 92%, Native American 4%, Hispanic or Latino 4% • Religion/belief– Not reported. • Disability– Not reported. • Long term health condition– Not reported. • Sexual orientation– Not reported. • Socioeconomic position– Not reported. • Type of abuse– EIFC: Sexual abuse 17%, physical abuse 24%, neglect 55%, emotional abuse 4% RFC: Sexual abuse 8%, physical abuse 4%, neglect 84%, emotional abuse 4%. Looked after or adopted status - Type of permanent placement EIFC: Reunification 48%, relative adoption 28%, nonrelative adoption 24%. Type of permanent placement RFC: Reunification 68%, relative adoption 20%, nonrelative adoption 12%. • Looked after or adopted status– Not reported. • Unaccompanied asylum seeking, refugee or trafficked children– Not reported.	When entered as a variable, membership of the intervention versus control condition did not significantly predict rates of failed placement (beta=-0.10, p=0.26). However, the interaction of condition x placements prior to study was significant (beta =-1.82, p=0.05), suggesting that 'the number of foster placements prior to the study was significantly related to failed permanent placements for children in RFC but not for children in EIFC' (p67). The interaction of condition x gender was marginally significant (beta=1.22, p=0.08). There was a higher failure rate for girls in the RFC, but not in the EIFC condition.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Sample size Comparison numbers -25 Intervention numbers -29 Sample size - 54 relating to permanent placement outcomes (overall study has 90 participants). **Intervention** **Describe intervention** Delivered via a team approach to the child, foster care provider and 'permanent placement resource' (p65) (birth parents, adoptive relatives or nonrelatives) to provide foster parents with intensive prior training, and ongoing consultation and support. **Delivered by** Clinicians with bachelor's and master's degrees and a licensed psychologist as clinical supervisor. **Delivered to** Before foster placement, foster parents given intensive training. After placement they receive support through daily telephone contacts, a weekly foster parent support group and 24-hour on-call crisis support. When a child is entering a permanent placement, the birth parents or adopters are trained in the same skills as the foster parents to support transition. Children receive services from a		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	behavioural specialist and attend weekly therapeutic playgroup sessions. **Duration, frequency, intensity, etc.** Typically 6 to 9 months, including transition to permanent placement. See above for frequency of contact. **Key components and objectives of intervention** Key features: - Following a developmental framework, characterising challenges faced by foster preschoolers as delayed development rather than strictly as emotional or behavioural problems - Encouraging prosocial behaviour in the child - Setting consistent limits to address disruptive behaviour - Close supervision of child - Development of a predictable daily routine. **Content/session titles** Not reported. **Location/place of delivery** Not reported. **Describe comparison intervention** Regular foster care in which children placed with foster carers and receive services as required. These can include: - individual mental health therapy - medical/dental treatment -		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	screening and referral for services for developmental delay. When child goes to permanent placement, services might include - social service support - substance abuse/mental health support for birth parents - parent training. **Outcomes measured** Service outcomes. Failure of permanent placement Number of foster care placements Time in foster care before a permanent placement.		

6. Forrester D, Holland S, Williams A et al (2014) Helping families where parents misuse drugs or alcohol? A mixed methods comparative evaluation of an intensive family preservation service. Child & Family Social Work 21(1) 65-75

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim Study aim: 1. To evaluate 'the impact of Option 2 on parental substance misuse and welfare, family functioning and child well-being' (p67). 2. To explore how 'parents view the service and its impact on their welfare' (p67). **Methodology** Mixed methods. A	**Participants** • Children and young people - 84 children, 34 parents or step-parents from 27 families. • Caregivers and families - 84 children, 34 parents or step-parents from 27 families. **Sample characteristics** • Age - Parents: age range 18->30 years Children: mean age: 9 years. • Sex - Most respondents were mothers (87%) Children: 59% males. • Ethnicity - All parents: most White	**Q15** Effect sizes **Effect sizes** **Quality of parenting and parent-child relationships** At follow-up Family functioning (FES score)(experienced poor functioning)(Table 2) Option 2: 7 (50%) vs control: 6 (60%); OR 1.5 (CI 0.29–7.75) Expressiveness scale (high better) Option 2:12.1 vs control:12.7 (ns) Cohesion scale (high better) Option 2: 11.6 vs control: 10.3 (mean difference = −1.27, 95% CI: −2.30 to −0.24) (p<0.01) Conflict scale (low is better) Option 2; 13.4 vs control: 14.0 (ns) Overall (high better) Option 2:10.3 vs control: 9.0 (ns) [At baseline, Moderate to high functioning) Option 2: 7 (50%) vs control: 4	**Overall assessment of external validity:** **++** **Overall assessment of internal validity:** **-** **Overall assessment of internal validity:** **-** **Overall validity rating:** **-** Unclear if the groups in

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
quasi-experimental study (comparative cohort) (quantitative) with interviews of parents (qualitative) to assess their views about the intervention. **Country** UK. **Source of funding** Voluntary/charity - funded by Alcohol Research UK.	British. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Child protection concerns with parents who misuse drugs or alcohol. Most of the families involved alcohol use issues (59%), a significant minority involved drug problems (44%). • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children – N.A. **Sample size** Comparison numbers - 12 families (28 children) in the comparison group. Intervention numbers - 15 families (46 children) received Option 2 Sample size - 1. Comparative cohort studies: 84 children, 34 parents or step-parents (quantitative data provided by 31) from 27 families. 2. Interview: Qualitative data from one or both parents for 26 families. Interviews were also carried out with children in five families (this is	(40%); OR 1.00 (at baseline)] Overall, families had more cohesion. **Children and young people's health and wellbeing outcomes** At follow-up, Child behaviour (SDQ score) (experienced some or high needs) (Table 2): Option 2: 6 (46%) v control; 3 (43%), OR 1.14 (CI 0.18–7.28) [Child behaviour (SDQ score) (experienced low needs at baseline): Option 2: 7 (54%) vs control: 4 (57%) OR 1.00 (Baseline)] There was no significant difference in the welfare of the children in the two groups. Child permanently moved (Table 4) Option 2: 17% v control: 41% (t=−2.27, p= 0.03) Entering care: Option 2: 8% vs control: 44% (t= −3.73, p=0.001) In permanent care: Option 2: 0% vs control: 38% (t=−4.31, p <0.001) Overall, Option 2 children less likely to enter care. **Caregiver/parent health and wellbeing outcomes** At follow-up, Reduction in parental substance: Option 2 : 17 (94%) vs control: 7 (58%); Odds ratio [OR] 12.14 (CI 1.19–123.62, p<0.05) (very wide CI due to small sample size)(Table 2) [Option 2: 1 (6%) vs control; 5 (42%); OR 1.00 (at baseline)] Parent's psychological distress (GHQ-12 score) experienced: Option 2: 19 (61%) vs control 11 (85%); OR 0.15 (CI 0.03–0.85, p<0.05) [Option 2; 10 (56%) vs control: 2 (15%); OR 1.00 (at baseline)] Overall, parents in Option 2 were more likely to have reduced their alcohol or drug misuse, they were less stressed and at risk of psychological problems. Other Detailed analyses of 'autobiographical narratives' about large parts of the lives of the participants were categorized into three groups: 1. Stories of change (n	the quasi-experimental cohort study are comparable.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	reported elsewhere, in another study). **Describe intervention** Option 2: "Based on the 'Homebuilders' model used and evaluated in the USA. It is based on a crisis intervention model, providing relatively brief (6 weeks) but very intensive input (workers have one primary case at any time and are available at all times).The communication style combines motivational interviewing and solution-focused approaches with family-orientated activities designed to build on strengths and core values held by parents. A full description is provided in Hamer 2005" (p67). **Delivered by** See under 'Describe intervention'. **Delivered to** Parents who misuse of drugs or alcohol. **Duration, frequency, intensity, etc.** See under 'Describe intervention'. **Key components and objectives of intervention** See under 'Describe intervention'. **Content/session titles** Option 2.	= 10) – in which there had been clear changes for the better such as reduced substance misuse and/or violent partners leaving [good outcomes]. 2. Stories of struggle (n = 7) – involved change, often partial and characterized by fluctuation between better times and increased difficulties [mixed outcomes] (p70). 3. Tales of troubles (n = 9) – 'descriptions of multiple and serious problems going on for years' (p70), 'with little reduction of drug or alcohol problems [poor outcomes]' (p70). The impact of Option 2 on these 3 groups: Option 2 was involved in: • 3 of 9 families with poor outcomes (33%) • 4 of 7 with mixed outcomes (57%) • 7 of 10 good outcomes (70%) suggesting that Option 2 was having a positive impact. Views and experiences of Option 2 users (see under 'narrative findings' qualitative and views and experiences).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Location/place of delivery Not reported. **Describe comparison intervention** Control: families referred when the service was full received basic information but do not receive a service (i.e. no O2). **Outcomes measured** Quality of parenting and parent-child relationships. Family functioning. Children and young people's health and wellbeing outcomes. Child behaviour. Caregiver/parent health and wellbeing outcomes. Reduction in parental substance misuse. Parent's psychological distress.		

7. Goldman Fraser J, Lloyd SW, Murphy RA et al. (2013) Child exposure to trauma: Comparative effectiveness of interventions addressing maltreatment. Comparative Effectiveness Review 89: 1–161

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim – 1. To evaluate the comparative efficacy and effectiveness of psychosocial and phar-	**Participants** Children and young people – 1. Aged 0–14 years of age, exposed to maltreatment (defined as child abuse [acts of commission: words or	**Effect sizes** Qualitative synthesis; a quantitative meta-analysis was not conducted due to the diversity in interventions, comparators, and outcomes measured.	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
macological interventions that address child well-being and/or promote positive child welfare outcomes (safety, placement stability, and permanency) for maltreated children ages birth to 14 years. 2. To assess the comparative effectiveness of interventions (a) with different treatment characteristics, (b) for child and caregiver subgroups, and (c) for engaging and retaining children and/or caregivers in treatment; 3. To assess harms associated with interventions for this population. **Methodology** Systematic review. – used review methods described in AHRQ's Methods Guide for Effectiveness and Comparative Effectiveness Reviews. data analysed in the form of a qualitative synthesis. A quantitative meta-analysis	overt actions that cause harm, potential harm, or threat of harm to a child] and child neglect [acts of omission: failure to provide for a child's basic physical, emotional, or educational needs or to protect a child from harm or potential harm]). 2. Children of the same ages involved with the child welfare system (including foster care), and caregivers of maltreated children when they were the target of an intervention. 3. Children with known CPS involvement. **Caregivers and families** – Also included primary caregiver(s) caregivers of maltreated children when they were the target of an intervention. **Sample characteristics** Age – Children 0–14 years of age; age of caregivers differed in different studies/interventions. Sex – Children and care givers (both male and female), details differed for different interventions. Ethnicity – Reported in some interventions: children and caregivers: Caucasians, African Americans, Hispanic/Latino, multiethnic populations. Religion/belief	Question 1: Comparative Effectiveness of Interventions for Improving Child Well-Being Outcomes and Improving Child Welfare Outcomes. Question 2. Comparative Effectiveness of Interventions for Improving Child Welfare Outcomes (safety: maltreatment recurrence). PARENTING INTERVENTIONS (Total 10 RCTs) 1. Parenting interventions: Attachment and Biobehavioral Catch Up (ABC) (3 RCTs, N=261). **Child wellbeing outcomes** **Mental and behaviour health** Compared with an active control (a home-based intervention focused on children's cognitive and linguistic development derived from the Abecedarian early intervention program) (1 RCT, n=120), children whose biological parents participated in ABC exhibited significantly less negative emotionality (effect size not reported; $p<0.05$). (strength of evidence [SOE] low). No significant difference in efficacy was found for ABC for parent report of child behavioral problems. (SOE low). Compared with a wait-list control (1 RCT, n=58), foster parents who participated in ABC reported significantly greater improvement in child internalising (partial eta squared =0.436; p=0.01) and externalising behavior (partial eta squared=0.511; p=0.001) (SOE low). **Healthy caregiver child relationship** Compared with active control (2 RCTs, n=213), children whose caregivers (foster or biological) participated in ABC exhibited significantly more positive attachment behaviours (fewer report of avoidant attachment behaviour, p=0.030; Decreased proportion with	+ **Overall score:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
was not performed due to issues of heterogeneity, insufficient numbers of similar studies, and poor outcome reporting. **Country** Range of countries – USA, UK, Canada. **Source of funding** Government – The Agency for Healthcare Research and Quality (AHRQ), Evidence-based Practice Centers (EPCs), USA.	– not reported. Disability – not reported. Long term health condition – not reported. Sexual orientation – not reported. Socioeconomic position – not reported. Type of abuse – included studies/interventions for sexually abused children. These will be presented under Q16. Looked after or adopted status – This including children who remain in the care of their biological parent and those in out-of-home care (e.g., foster care, kinship care, group home care). Unaccompanied asylum seeking, refugee or trafficked children – not reported. **Sample size** Systematic reviews: number of studies – qualitative synthesis of 11 trials on physical abuse and neglect. Studies with a high risk of bias are excluded by the authors in the results. **Intervention category** Parenting intervention – A. Parenting interventions: 1. Attachment and Biobehavioral Catch Up vs	disorganized attachment, effect size not reported; p<0.01; Increased proportion with secure attachment, effect size not reported; p<0.05) (SOE low). Compared with a wait-list control (1 RCT, n=58), foster parents who participated in ABC had greater improvement in parent attitudes (Improvements in self-reported risk factors for child abuse, partial eta squared=0.791; p=0.001; and greater reductions in parent stress (partial eta squared=0.59; p=0.01) (SOE low). **Healthy development** Compared with an active control, children whose foster parents participated in ABC exhibited higher levels of cognitive functioning (Cognitive flexibility, effect size not reported; p=.008; Theory of mind, effect size not reported; p=.01) (SOE low). NB. Parenting interventions: Attachment and Biobehavioral Catch Up: No studies assessed child welfare outcomes (safety: maltreatment recurrence). **2. Parenting interventions: Attachment-Based Intervention** One RCT, n=79. **Child wellbeing outcomes** **Mental and behaviour health** No significant differences in efficacy of the intervention on child internalising or externalising behaviour (SOE insufficient). **Healthy caregiver child relationship**	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	active control vs waiting list; 2. Attachment- Based Intervention vs. usual care; 3. Child-Parent Psychotherapy vs active control vs usual care; 4. Incredible Years Adaptation vs. usual care; 5. Keeping Foster and Kinship Parents Trained and Supported vs usual care; 6. Nurse-Home Visitation Intervention vs usual care; 7. PCIT Adaptation Package vs PCIT Adaptation Package Enhanced vs Usual care; 8. PCIT Adaptation Package Enhanced vs usual care; 9. Safe care vs. usual care; Parent-Child Interaction therapy – see details in Parenting interventions Other. Intervention **Describe intervention** – A. Parenting interventions (10 trials) 1. Attachment and Biobehavioral Catch-up (ABC), a caregiver-directed approach to guide and support nurturing, sensitive care that promotes child regulatory capabilities and attachment formation. Manualised curriculum, allows flexibility in responding to current issues; videotaped parent–child interactions used to illuminate child cues and strengths in the relationship. Duration/intensity: 10, 1-hr weekly home visits with caregiver and child together. Target population: children aged from	Compared with usual care, Participants in the Attachment-based Intervention demonstrated significant improvements in maternal sensitivity (d=0.47; p<0.05) and secure attachment behaviour (r=0.36; p<0.05), changes from disorganized to organized attachments (r=0.37; p<0.05) (SOE low). **Healthy development** No data. NB. Attachment-Based Intervention: No studies assessed child welfare outcomes (safety: maltreatment recurrence). **3. Parenting interventions: Child-Parent Psychotherapy** Two RCTs, n=224. **Child wellbeing outcomes** **Mental and behaviour health** No data **Healthy caregiver-child relationship** Compared with an active control, preschool-age children who participated in CPP reported significantly fewer negative attachment representations (greater decline in negative self-representations, effect size not reported; p=0.01); however, for younger children, there were no significant differences in efficacy of the intervention on secure attachment behaviour (SOE insufficient). When compared with usual care, infants who participated in CPP demonstrated significantly greater improvements in secure attachment behaviour (higher	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	birth to 5 years and their primary caregivers. 2. Attachment-based Intervention, a caregiver-directed approach, loosely derived from ABC and other attachment-focused interventions, to guide and support maternal sensitivity to child cues and secure attachment. Employs individualized parent-child interaction support, video feedback, and discussion of attachment and emotion regulation-related themes. Duration/intensity: 8, 1.5-hr weekly home visits with caregiver and child together. Target population: children aged 1–5 years and their caregivers. 3. Child-Parent Psychotherapy (CPP), a relationship-based dyadic psychotherapy, with focus on supporting formation of and repairing the caregiver-child attachment relationship; home- or clinic-based. Employs the parent-child relationship as the 'port of entry' for therapeutic work. Duration/intensity: 50, approximately 1-h weekly visits with child and caregiver together. Target population: children aged 1–5 years and their primary caregivers. 4. Incredible Years Adaptation, a caregiver-directed approach adapted for use with foster and biological parent pairs to address placement issues (e.g., safety; attachment and loss); supplemented with a co-parenting	rates of secure attachment, h=1.16 to 1.39; p=0.01; Higher rates of becoming securely attached, h=1.16 to 1.34; p=0.01; Lower rates of stable disorganized attachment, h=0.64 to 0.83; p=0.025); and preschool-age children reported significantly fewer negative attachment representations (Greater decline in negative self-representations, effect size not reported; p=0.01) (low SOE). **Healthy development** No data. NB. Parenting interventions: Child-Parent Psychotherapy: No studies assessed child welfare outcomes (safety: maltreatment recurrence). **4. Parenting interventions: Incredible Years Adaptation** One RCT, n=128 carers and 64 children. **Child wellbeing outcomes** **Mental and behavioural health** No significant differences in efficacy of the IY Adaptation with parents (biological and foster) on child internalizing or externalizing problems (Caregiver perception of child behavioural problems; Caregiver perception of child behavioural and conduct problems; Teacher report of disruptive classroom behaviours (SOE insufficient). **Healthy caregiver-child relationship** Compared with usual care, parents (biological and foster) who participated in the IY Adaptation reported a significant increase in the endorsement/use of positive	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	component designed to support a positive, nonconflicted relationship between caregivers and increase caregiver sensitivity. Duration/intensity:12, 2-hr weekly parent group sessions for biological-foster parent pairs, supplemented with weekly sessions (duration not specified) with individual families (biological and foster parent pair and target child). Target population: caregivers of children aged 3–10 years. 5. Keeping Foster Parents Trained and Supported (KEEP), a parent training to increase foster/kin caregivers' use of positive discipline strategies. Delivered by paraprofessionals; employs role plays, videotapes, homework practice. Duration/intensity: 16, 1.5-hr weekly parent group sessions, with 15-min didactic presentations by facilitators then group discussion related to primary curriculum concepts. Target population: caregivers of children aged 5–12 years. 6. Nurse Home Visitation Intervention A caregiver-directed approach offering intensive family support, parent education, and referrals to health and social services (derived from Olds et al. 1997 home visiting preventive intervention (authors developed their own manual). Employs mutual problem identification, goal setting, and problem-solving strategies; supporting positive parent-child	parenting practices (Greater reporting of positive discipline strategies, d=0.40 to 0.59; p=0.01 at T2; Greater reporting of setting clear expectations for child, d=0.54; p<0.05 at T2 (SOE low). **Healthy development** No data. NB. Parenting interventions: Incredible Years Adaptation: No studies assessed child welfare outcomes (safety: maltreatment recurrence). **5. Parenting interventions: Keeping Foster and Kinship Parents Trained and Supported (KEEP)** One RCT, n=700. **Child wellbeing outcomes** **Mental and behavioural health** Compared with usual care, Participants in KEEP reported significantly greater improvement in child externalising behaviour (Improvement in problem behaviours at end point, d=0.26, significant but p value not reported) (SOE moderate). **Healthy caregiver-child relationship** Compared with usual care, Participants in KEEP reported significantly increased use of positive discipline practices (Increased proportion positive reinforcement at endpoint significant, d=0.29, but p value not reported) (SOE moderate). **Healthy development** No data.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	interaction. Duration/intensity: 6 months of 1.5-hr weekly home visits with parent, then visits every 2 weeks for 6 months, then monthly visits for 12 months. Target population: caregivers of children aged from birth to 13 years. 7. SafeCare A home-based multifaceted parent services to prevent and treat child abuse and neglect, formerly known as Project 12-Ways. The services address parent-child or parent-infant interaction, parental stress, and home safety risks including behavior management, problem solving, infant and child health and nutrition, and social support. Duration/intensity: Home visits at least weekly for 6 months (duration not specified) Target population: Children ages 0 to 12 years. 8. Parent-Child Interaction Therapy Adaptation A standard parent-child interaction therapy (PCIT), based on social learning and attachment theory, adapted for abusive or neglectful parents. It includes a 3-phases motivational intervention orientation: (1) motivational intervention (orientation phase); (2) child-directed interaction phase during which parents develop child-centred interaction skills; (3) parent-directed interaction phase during which effective discipline skills are the focus. It uses live parent-child skills practice/rehearsal, with live coaching	NB. Keeping Foster and Kinship Parents Trained and Supported (KEEP): report on placement outcomes only. **6. Parenting interventions: Nurse Home Visitation Intervention (NHV)** One RCT, n=163. **Mental and behavioural health** No significant differences in efficacy of the intervention on child internalizing or externalising behaviour (SOE insufficient). **Healthy caregiver-child relationship** No significant differences in efficacy of the intervention parent attitudes, parenting practices associated with child abuse, family functioning, or the home environment (SOE insufficient). **Healthy development** No data. NB. Parenting interventions: Nurse Home Visitation Intervention: Child welfare outcomes (Safety: maltreatment recurrence). Compared with usual care, there was no significant difference in efficacy of the intervention on maltreatment recurrence based on Child Protective Services (CPS) records, whereas hospital records showed significantly higher rates of recidivism for the NHV condition compared with usual care (23.6% vs. 10.8%, difference 12.8% [95% CI, 1.4 to 24.1]). There was a borderline significant difference (effect size not reported;	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	by the therapist (immediate feedback from therapist from observation room to parent via wireless earphone). Coaching is driven by behavioural principles such as modelling, reinforcement, and selective attending to shape parents' behaviours. Duration/intensity: Motivational intervention: 6 clinic-based parent group sessions/therapeutic sessions: 12 to 14 approximately 1-hour clinic-based individual sessions with parent and child together. Target population: Children ages 4 to 12 years. 9. Videotape Intervention (for sexually abused children) Informed by social learning theory to increase supportive maternal behaviours following sexual abuse of a child and the child's subsequent medical evaluation. The videotape provides specific information about short- and long-term psychological and behavioural effects seen in sexually abused children, reactions of parents, and importance of how parent respond to children; suggested responses presented as 'BRAVE To Tell' representing five specific supportive behavioural approaches for interacting with child. Duration/intensity: 22-minute videotape presented to parents during child's forensic examination Target population: Children ages 4 to 12 years.	p=0.058) between the intervention and comparison groups in the severity of neglect incidents, favouring the intervention group.(SOE insufficient). **7. Parenting interventions: Parent-Child Interaction Therapy Adaptation Package (PCIT-AP)** Two RCTs, n=263. **Child wellbeing outcomes** No data. **Child welfare outcomes** Safety: compared with usual care, PCIT-AP was reported to reduce child maltreatment recidivism (reports to the child welfare system) (effect size not reported; p=0.02) (1 RCT, N=110) (SOE low). Compared with PCIT-AP, an enhanced version of the intervention that provided individualised services and home visits showed no significant difference in efficacy on recidivism (1 RCT, n= 75) (insufficient SOE). Compared with the community standard parenting program combined with the experimental self-motivational orientation, PCIT–AP resulted in significantly reduced recidivism (recurrence of maltreatment, effect size not reported; p=0.05) (1 RCT, n=153) (low SOE). **8. Parenting Intervention: SafeCare** One RCT, n=2175. **Child wellbeing outcomes** No data. **Child welfare outcomes**	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Delivered by – see 'Describe interventions'. **Delivered to** – see 'Describe interventions'. **Duration, frequency, intensity** – see 'Describe interventions'. **Key components and objectives of intervention** – see 'Describe interventions'. **Content/session titles** – see 'Describe interventions'. **Location/place of delivery** – see 'Describe interventions'. Settings varied with interventions: outpatient and inpatient mental healthcare settings; schools, community-based providers, shelters, prison; home-based and out-of-home care (e.g., foster or kin care, residential treatment). **Describe comparison intervention** –Usual care, active control (interventions representative of conventional practices in the field [family, child-centred, or supportive group therapy] or modified version of the intervention model, inactive control.	Safety: Compared with usual care, SafeCare resulted in significantly reduced child maltreatment recidivism (reports to the child welfare system) (Hazard ratio [HR]=0.83 [95% CI, 0.70 to 0.98] for the full population; HR=0.74 [95% CI, 0.58 to 0.95] for the preschool sub-population (SOE moderate). 9. Video tape intervention (not data extracted – population children who have been sexually abused, will be data extracted for Q16). **TRAUMA-FOCUSED INTERVENTIONS (1 RCT)** **10. Trauma-Focused Treatments: Combined Parent-Child Cognitive Behavioural Therapy** (CPCCBT) One RCT, n=75. **Child wellbeing outcomes** **Mental and behavioural health** Compared with an active control, Participants in CPCCBT had a significantly greater reduction in trauma symptoms (Parent and child report of trauma symptoms, d=0.61; p<0.05); however, there was no significant difference in efficacy of the intervention on child internalizing or externalizing behaviour problems (SOE low). **Healthy caregiver-child relationship** Compared with an active control, parents in CPCCBT reported significantly greater increases in positive parenting practices (parent report of positive parenting, d=0.59; p<0.05; parent report of reduction in use of corporal punishment, d=0.57; p<0.05). Based on child report, there were no significant differences in efficacy	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Outcomes measured Incidence of abuse and neglect – Safety (i.e., maltreatment recurrence) Risk of abuse and neglect – Self-reported risk factors for child abuse Quality of parenting and parent-child relationships – Healthy caregiver-child relationship outcomes Children and young people's health and wellbeing outcomes – Children's mental and behavioural health Caregiver/parent health and wellbeing outcomes – parental stress.	of the intervention on positive parenting practices or use of corporal punishment (SOE insufficient). **Healthy development** No data. NB. Trauma-Focused Treatments: CPCCBT: No studies assessed child welfare outcomes (Safety: maltreatment recurrence). 11. Trauma-Focused Treatments: Eye Movement Desensitization and Reprocessing (EMDR) (1 RCT, n=14) (sexually abused children) To DE for Q16. 12. Trauma-Focused Treatments: Group Psychotherapy for Sexually Abused Girls (1 RCT, n=71) (sexually abused children) To DE for Q16. 13. Trauma-Focused Treatments: Group Treatment Program for Sexual Abuse (1 non-RCT, n=30) (sexually abused children) To DE for Q16. 14. Trauma-Focused Treatments: Trauma-focused cognitive behavioral therapy (3 RCTs, n=359) (sexually abused children) To DE for Q16. **Question 3. Comparative Effectiveness of Interventions With Different Characteristics: theoretical orientation (3 RCTs)** Attachment-based approach versus didactic approach Attachment based intervention was found to show benefit in child mental and behavioural health, caregiver-child relationship and developmental outcomes	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		(no effect size reported), when compared with didactic approach (low SOE). Cognitive behavioural versus psychodynamic Cognitive behavioural intervention was found to show benefit in child mental and behavioural health (d=0.33 to 0.70) and caregiver-child relationship (d=0.38 to 0.57) outcomes, when compared with the psychodynamic approach (low SOE). **Question 4. Comparison of Intervention Effectiveness for Improving Child Well-Being or Child Welfare Outcomes in Population Subgroups** **Question 4a. Child Well-Being and Child Welfare Outcomes in Child Subgroups** Parenting interventions: Early childhood (ages 0–5 years) (7 RCTs): Attachment and Biobehavioral Catch Up (ABC) resulted in improvements in child mental and behavioural health and caregiver-child relationship outcomes compared with an active control (2 RCTs), and with waiting list control (1 RCT) (SOE low). Attachment-Based Intervention was reported to improve caregiver-child relationship outcomes compared with usual care (1 RCT) (SOE low). Child-Parent Psychotherapy (CPP) was reported to result in improved caregiver-child relationship outcomes compared with usual care (2 RCTs) (SOE low). SafeCare resulted in significantly reduced child recidivism (re-reports to child welfare) compared with usual care; the benefits of SafeCare were strongest for pre-	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		school-age children compared with the full study population, which included children up to 12 years of age (1 RCT) (SOE moderate). Subgroup by type of maltreatment: Neglect Parenting interventions: SafeCare intervention resulted in significantly reduced child recidivism (re-reports to child welfare) compared with usual care. More than a third of the children in each study group had previous histories of physical abuse and of sexual abuse. The benefits of SafeCare were strongest for preschool-age children compared with the full study population, which included children up to 12 years of age (1 RCT) (SOE moderate). Subgroup by type of maltreatment: physical abuse (2 RCTs) Parenting Interventions: Parent-Child Interaction Therapy Adaptation Package (PCIT-AP) including a motivational interviewing orientation, found significantly reduced child recidivism in favour of the intervention (1 RCT) (SOE low). Trauma-Focused Treatments: Combined Parent-Child Cognitive Behavioural Therapy (CPC-CBT), designed specifically to treat children exposed to physical abuse, found greater improvements in child mental and behavioural health among children in the intervention group compared with an inactive control (1 RCT) (SOE low). **Question 4b. Child Welfare and Child Well-Being Outcomes in Caregiver Subgroups** Maltreating parents (7 RCTs and 1 cohort study):	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Parenting interventions: Attachment and Biobehavioural Catch Up (ABC): The study found significantly greater improvements in child mental and behavioural health and caregiver child relationship outcomes (partial eta squared 0.436 to 0.511) compared with a waiting list. (1 RCT) (SOE low). Parenting interventions: Attachment-Based Intervention: The study found improved caregiver-child relationship outcomes (d=0.47, r=0.36) compared with usual care (1 RCT) (SOE low). Parenting interventions: Child-Parent Psychotherapy (CPP) resulted in improved caregiver-child relationship outcomes (h=0.64 to 1.34) compared with usual care (2 RCTs) (SOE low). Parenting interventions: Parent-Child Interaction Therapy Adaptation Package (PCIT-AP): this study found significantly reduced child recidivism in favour of the intervention (1 RCT); a trend towards this effect was found in a subsequent effectiveness trial which targeted parents referred for services by child welfare for neglect and/or physical abuse (1 RCT)(SOE low). Parenting interventions: SafeCare intervention resulted in significantly reduced child recidivism (re-reports to child welfare; HR 0.74 to 0.83) compared with usual care. The benefits of SafeCare were strongest for preschool-age children compared with the full study population, which included children up to 12 years of age (1 RCT) (SOE moderate). Parenting interventions: Incredible Years Adaptation (IYA) for Neglecting parents: The study found significantly greater improvements in caregiver child relationship outcomes (d=0.40 or 0.59) compared with usual care.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Trauma-Focused Treatments: Combined Parent-Child Cognitive Behavioural Therapy (CPC-CBT): This study found significantly greater improvements in child mental and behavioural health outcomes (d=0.61) compared with an active control; but the short-term outcomes faded by 3-month postintervention (1 RCT) (SOE low). Foster parents (3 RCTs) Parenting interventions: Attachment and Biobehavioural Catch Up (ABC): This study found significantly greater improvements in child mental and behavioural health outcomes, caregiver child relationship and child developmental outcomes when compared with active control (SOE low). Parenting interventions: Attachment and Biobehavioural Catch Up (ABC): This study found significantly greater improvements in child mental and behavioural health and caregiver child relationship outcomes when compared inactive control (SOE low). Parenting interventions: Keeping Foster and Kinship Parents Trained and Supported (KEEP): This study found significantly greater improvements in child mental and behavioural health (d=0.26) and caregiver child relationship outcomes (d=0.29) when compared with inactive control (SOE moderate). **Question 5. Comparative Effectiveness of Interventions With Children Exposed to Maltreatment for Engaging Children and/or Caregivers in Treatment** (1 RCT, n=153) Treatment engagement: Compared with parents who participated in a standard orientation, maltreating parents who participated in the motivational intervention	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		(MI) orientation reported greater readiness for change (d=0.33, p<0.01) and other positive self-motivational outcomes (Increased readiness to change, p<0.05; Better attitude to the program, p< 0.05)(SOE moderate). Treatment retention: Compared with parents who participated in a standard orientation combined with PCIT, maltreating parents who participated in PCIT combined with the MI orientation had higher treatment completion rates (Higher percentage of treatment completers, p=0.01 to 0.05) (SOE moderate). **Question 6. Adverse Events Associated With Interventions for Children Exposed to Maltreatment (Data available for sexually abused children) To DE for Q16.** No eligible studies on pharmacotherapy was identified by the review. **Narrative findings** This systematic review (Goldman Fraser 2013 +) (15 RCTs and 1 cohort studies) assessed the effectiveness of psychosocial interventions in children aged 0-14 years who are exposed to maltreatment (physical abuse and neglect). The evidence base for effective interventions is limited with predominantly low level of evidence derived from results of single trials evaluating different outcomes. Firm conclusion of the review cannot be drawn. PARENTING INTERVENTIONS: **1.Attachment and Biobehavioral Catch Up programme (ABC)(3 RCTs, evidence level low)**	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Compared with active control, children whose biological parents participated in ABC showed significantly less negative emotionality (1 RCT), more positive attachment behaviours, decreased proportion with disorganized attachment and increased proportion with secure attachment. These children also exhibited higher levels of cognitive functioning. Compared with a wait-list control, foster parents who participated in ABC reported significantly greater improvement in child internalizing and externalizing behaviour (small to medium effect size, partial eta squared 0.44 to 0.51), and had greater improvement in parent attitudes (large effect size, partial eta squared 0.79) and greater reductions in parental stress (medium effect size, partial eta squared 0.59). **2.Attachment-Based Intervention (1 RCT, evidence level low)** Compared with usual care, participants in the Attachment-based Intervention reported significant improvements in maternal sensitivity (small to medium effect size, d=0.7) and secure attachment behaviour (small to medium effect size, r=0.36) and improved organized attachments (small to medium effect size, r=0.37). **3. Child-Parent Psychotherapy (CPP)(2 RCTs, evidence level low/insufficient)** Compared with an active control, preschool-age children who participated in CPP reported significantly fewer negative attachment representations but no significant differences in efficacy of the intervention on secure attachment behaviour.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Compared with usual care, infants in CPP demonstrated significantly greater improvements in secure attachment behaviour (small effect size, h=1.6); lower rate of disorganised attachment (medium to large effect size, h=0.64 to 0.83) and preschool-age children reported significantly fewer negative attachment representations. **4. Incredible Years Adaptation (IYA)(1 RCT, evidence level low)** Compared with usual care, there was no significant difference in parent-reported child internalising or externalizing problems after IY Adaptation Intervention. Compared with usual care, parents (biological and foster) who participated in the IY Adaptation reported a significant increase in the endorsement/use of positive parenting practices (small to medium effect size, d=0.40 to 0.59). **5.Keeping Foster and Kinship Parents Trained and Supported (KEEP)** (1 RCT, evidence level moderate) Compared with usual care, participants in KEEP reported significantly greater improvement in child externalizing behaviour (small effect size, d=0.26). Compared with usual care, participants in KEEP reported significantly increased use of positive discipline practices (small effect size, d=0.29). **6. Nurse Home Visitation Intervention (NHV)** (1 RCT, evidence level insufficient) Compared with usual care, there was no significant difference in the NHV group on child internalising or externalizing behaviour, on parent attitudes, parenting	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		practices associated with child abuse, family functioning, or the home environment. Compared with usual care, there was no significant difference in the NHV intervention on maltreatment recurrence based on Child Protective Services (CPS) records whereas hospital records showed significantly higher rates of recidivism for the NHV condition (mean difference 10.8%). **7. Parent-Child Interaction Therapy Adaptation Package (PCIT-AP)** (2 RCTs, evidence low/insufficient) Compared with usual care, PCIT-AP was reported to reduce child maltreatment recidivism (reports to the child welfare system). Compared with the community standard parenting program combined with the experimental self-motivational orientation, PCIT–AP resulted in significantly reduced recidivism. Compared with PCIT-AP, an enhanced version of the intervention that provided individualised services and home visits showed no significant difference in efficacy on recidivism (recurrence of maltreatment). **8. SafeCare** (1 RCT, evidence level moderate) Compared with usual care, SafeCare resulted in significantly reduced child maltreatment recidivism (reports to the child welfare system) for the preschool subpopulation (Hazards ratio 0.74, 95% CI 0.58 to 0.95). TRAUMA-FOCUSED INTERVENTIONS **Combined Parent-Child Cognitive Behavioral Therapy (CPCCBT)** (1 RCT, evidence level low/insufficient)	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Compared with an active control, participants in CPCCBT had a significantly greater reduction in trauma symptoms (medium effect size, d=0.59) but no significant difference in efficacy of the intervention on child internalizing or externalizing behaviour problems. Parents in CPCCBT also reported significantly greater increases in positive parenting practices (medium effect size, d=0.57 to 0.59). However, there were no significant differences in efficacy of the intervention on child-reported positive parenting practices or use of corporal punishment. ADVERSE OUTCOMES Available data relate only to interventions for sexually abused children – to be data extracted for Q16. COMPARATIVE EFFECTIVENESS OF INTERVENTIONS WITH DIFFERENT CHARACTERISTICS: THEORETICAL ORIENTATION (2 interventions assessed this subgroup). Compared with didactic interventions (non-relationship-based), families who participated in attachment-based intervention were more likely to achieve better outcomes in child mental and behavioural health, caregiver-child relationship and child healthy development. Compared with psychodynamic interventions, families who participated in cognitive behavioural interventions were more likely to achieve better outcomes in child mental and behavioural health and caregiver-child relationship.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		COMPARATIVE EFFECTIVENESS OF INTERVENTIONS WITH DIFFERENT CHARACTERISTICS: EARLY CHILDHOOD (0–5 YEARS AGE) (4 interventions assessed this age group) Compared with usual care or waiting list control, families with children aged 0–5 years (Early childhood) who participated in ABC, ABI, CPP and SafeCare were more likely to achieve better outcomes in child mental and behavioural health, caregiver-child relationship, child development and reduced maltreatment recurrence. COMPARATIVE EFFECTIVENESS OF INTERVENTIONS WITH DIFFERENT CHARACTERISTICS: NEGLECT ONLY (1 intervention assessed this subgroup) Compared with usual care, families offered SafeCare were more likely to achieve better child mental and behavioural health, caregiver-child relationship, child development and reduced maltreatment recurrence. COMPARATIVE EFFECTIVENESS OF INTERVENTIONS WITH DIFFERENT CHARACTERISTICS: PHYSICAL ABUSE ONLY (1 intervention assessed this subgroup) Compared with active control, families offered CPCCBT were more likely to achieve better child mental and behavioural health outcomes. COMPARATIVE EFFECTIVENESS OF INTERVENTIONS WITH DIFFERENT CHARACTERISTICS: MALTREATING PARENTS (7 interventions assessed this subgroup)	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Compared with usual care, maltreating parents offered ABI and CPP were more likely to achieve better care-giver-child relationship outcomes. Compared with usual care, maltreating parents offered PCIT-AP were more likely to achieve reduction in their maltreatment recurrence. Compared with active control, maltreating parents offered ABC were more likely to achieve better child mental and behavioural health and caregiver-child relationship outcomes. Compared with active control, maltreating parents offered CPCCBT were more likely to achieve better child mental and behavioural health outcomes. Compared with inactive control, maltreating parents offered IYA were more likely to achieve better child mental and behavioural health and caregiver-child relationship outcomes. Compared with usual care, maltreating parents offered SafeCare were more likely to achieve reduction in their maltreatment recurrence. COMPARATIVE EFFECTIVENESS OF INTERVENTIONS WITH DIFFERENT CHARACTERISTICS: FOSTER PARENTS (3 interventions assessed this subgroup) Compared with usual care, foster parents offered KEEP were more likely to achieve better child mental and behavioural health and caregiver-child relationship outcomes.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Compared with usual care, foster parents offered MTFC-P were more likely to achieve better child mental and behavioural health, caregiver-child relationship, child development outcomes. Compared with active control, foster parents offered ABC were more likely to achieve better child mental and behavioural health, caregiver-child relationship, child development outcomes. Compared with inactive control, foster parents offered ABC were more likely to achieve better child mental and behavioural health and child development outcomes. COMPARATIVE EFFECTIVENESS OF INTERVENTIONS FOR ENGAGING CHILDREN AND CAREGIVERS IN TREATMENT (1 intervention) Treatment engagement: Compared with standard approach, maltreating parents offered motivational intervention (MI) were more likely to report greater readiness to change and to complete interventions. Treatment retention; Compared with a standard approach combined with PCIT, maltreating parents offered PCIT combined with motivational intervention (MI) were more likely to complete treatment. No eligible studies on pharmacotherapy was identified by the review authors. LIMITATIONS No quantitative meta-analysis due to heterogeneity.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Results derived from single trials, many with small sample size. Firm conclusion of the review cannot be drawn. Substantive methodological limitations in included studies: small sample, size, poor reporting, unclear definitions, short follow-ups. Interventions of different intensity. Data from studies conducted in US, Canada and Romania, issues of generalisability due to differences in child welfare systems and health service systems. This review is rated ++ for internal validity and + external validity, overall score of + for validity.	

8. Graham-Bermann SA, Miller-Graff LE, Howell KH et al. (2015) An Efficacy Trial of an Intervention Program for Children Exposed to Intimate Partner Violence. Child Psychiatry and Human Development 46: 928–939

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim Study aim: 'To compare outcomes for 4-6 year old children randomly assigned to a program designed to address the effects of exposure to IPV with those allocated to a waitlist comparison condition' (p928). **Methodology** RCT including cluster.	**Participants** Children and young people - Children exposed to intimate partner violence Caregivers and families - Mothers who have experienced intimate partner violence. **Sample characteristics** • Age - Children aged 4 to 6, mean age 4.93 (sd=0.86). Mothers: mean age 31.9 (sd=7.19). • Sex - Children: Male 53%, female	**Effect sizes** Children and young people's health and wellbeing outcomes. Note: Findings reported in text do not match results in Table 2. Findings taken from ITT analysis reported in Table. Using multilevel regression, there were no significant effects for: - Treatment - Time 2 - Time 3 - Child sex. However, there was a significant treatment x time 3 interaction (beta=-0.475), $p<0.05$, although not a significant treatment x time 2 interaction (beta=-0.111, $p>0.05$). The discussion states that there was an improvement in internalising symptoms for girls only. However, this does not match the findings as reported in Table 2.	**Overall assessment of internal validity** + **Overall assessment of external validity** + **Overall validity score** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Country Not UK. USA. **Source of funding** Not reported.	47%. All caregivers were female. • Ethnicity - Children: Caucasian 38%, African American 37%, 'biracial' 20%, Latino/a 5% Mothers: Caucasian 48%, African American 37%, 'biracial' 8%, Latino/a 6%, Other 1%. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Mean monthly income $1414, (sd=$1549). • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** • Comparison numbers - n=62 • Intervention numbers - n=58 • Sample size - n=120 **Describe intervention** The Pre Kids Club (PKC) intervention has two components: PKC for children and the Moms' Empowerment Programme (MEP) for mothers. Based on the assumption that children may be		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	traumatised by witnessing IPV, and develop unhelpful attitudes and beliefs as a result. PKC involves discussing issues related to IPV in an age-appropriate way, using a training manual. MEP is designed to support mothers' social and emotional adjustment. As part of the intervention mothers are support to discuss: the impact of IPV on their child their mental health symptoms; normalise and reduce stress; provide support regarding parenting challenges. **Delivered by** Master's level social workers and graduate students in clinical psychology. **Duration, frequency, intensity, etc.** Ten-session intervention provided over five weeks. **Content/session titles** Sessions focus on topics related to intimate partner violence, including: attitudes and beliefs about violence; managing emotions; safety planning; conflict resolution. **Describe comparison intervention** Waitlist. **Outcomes measured**		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Children and young people's health and wellbeing outcomes. Child adjustment (internalising behaviour) measured using the Child Behavior Checklist (Achenbach 1991).		

9. Jouriles EN, McDonald R, Rosenfield D et al. (2010) Improving parenting in families referred for child maltreatment: A randomized controlled trial examining effects of Project Support. Journal of Family Psychology 24:328–38

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim – The study aimed to evaluate the effectiveness of Project Support, a home-based intervention targeting parenting and maternal distress, in comparison to services as usual, in a sample of families referred for child maltreatment. **Methodology** RCT – Participants were randomly allocated to one of two conditions: the intervention and services as usual. **Country** Not UK – US study.	**Participants** **Caregivers and families** – Participants were families in which allegations of physical abuse or neglect of a child aged 3–8 were substantiated by CPS, and in which it was determined that the child and family's interests would be best served by keeping the family intact, and the mother (or both parents) would participate in services. All primary caregivers in the study were mothers. **Sample characteristics** Age – The average age of mothers was 28.7 (SD=5.4) and the average age of children was 5.4 (SD=1.5). Sex – The exact number is not stated, however, it is gleaned from the information that the sample was entirely female.	**Effect sizes** **Incidence of abuse and neglect** – No significant difference in re-referrals to CPS between Project Support families versus comparison, but with medium effect size. Project Support 1/17 (5.9%) re-referrals, comparison 5/18 (27.7%), chi-squared (1)=2.95, p=0.086, effect size (phi)=0.29. **Quality of parenting and parent-child relationships** – 1. Self-reported inability to manage childbearing responsibilities, measured using the Parenting Locus of Control Scale (PLOC) (NB decrease in scores reflects improvements in parenting). Using hierarchical linear modelling the study found that scores for the Project Support group decreased (i.e. improved) to a significantly greater extent than for the comparison group b= 1.09, t(32)=2.58, p<.05, ES=1.02, 95% CI [0.29, 1.70]. 2. Self-reported harsh parenting behaviours, measured using the psychological aggression and minor assault subscales from the Revised Conflict Tactics Scales (CTS-R). The study found that harsh parenting	**Overall assessment of internal validity** + **Overall assessment of external validity** ++ **Overall validity score** + – Key limitations of the study are: Assessors not blind to participant condition; Comparison intervention was not consistent across all participants; Method for intent-to-treat analysis not reported.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Source of funding Other – The study was funded by grants from the Interagency Consortium on Violence Against Women and Violence Within the Family, and the Hogg Foundation for Mental Health.	Ethnicity – 23% of mothers reported their ethnicity as White, 47% as Black, 26% as Hispanic and 3% as other. Religion/belief – This was not stated. Disability – This was not stated. Long term health condition – This was not stated, although mothers experiencing serious mental health issues and substance misuse disorders were excluded from the study. Sexual orientation – This was not stated. Socioeconomic position – This was not stated, although the authors note that the intervention was augmented to address the circumstances of children in abusive, low-income families. Type of abuse – 63% had been referred to CPS for physical abuse, 25% for neglect and 12% for both. Looked after or adopted status – Participants included families that were still intact, but had been reported to CPS for child maltreatment. Unaccompanied asylum seeking, refugee or trafficked children – This is not stated.	scores also decreased to a significantly greater extent in the Parent Support group compared to the comparison group b=0.14, t(32)=2.26, p<.05, ES=0.86, 95% CI [0.15, 1.53]. 3. Staff observations of ineffective parenting Staff observations of ineffective parenting showed significantly greater improvement in the Parent Support group compared to the comparison group b= 0.38, t(32)= 2.22, p<.05, ES=0.96, 95% CI [0.24, 1.64]. **Caregiver/parent health and wellbeing outcomes** – 1. Maternal psychological distress, measured using the Symptom Checklist-90-Revised (SCL-90-R). There was no significant difference in changes to maternal psychological distress between Project Support and the comparison condition (data not provided). **Narrative findings - effectiveness** – Findings suggest that Project Support had meaningful effects on mothers' parenting. Specifically, mothers in the Project Support condition, compared to those in the comparison condition, showed greater reductions in perceived inability to manage their children's behaviour, self-reported harsh parenting and observed ineffective parenting. Improvements were most rapid during treatment and maintained during follow-up. The intervention had no statistically significant impact on subsequent referrals to Child Protective Services, although the authors note that this was large in absolute terms (5.9% of families compared to 27.7% of families). However, it should be noted that the effect size was small (0.29). The rate of improvement in mothers' psychological distress did not differ significantly across the two groups.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Sample size Comparison numbers – n=18. Intervention number – n=17. **Sample size** – A total of 35 families were included in the study. **Intervention** **Describe intervention** – Project Support is a home-based intervention involving two primary components: 1) teaching mothers child behaviour management skills and 2) providing instrumental and emotional support to mothers. The authors hypothesise that the primary mechanism for reducing maltreatment is the child behaviour management skills component. Mothers are taught skills with which to increase desirable, and decrease undesirable, child behaviours and facilitate a positive and warm relationships. This is taught through direct instruction, practice and feedback. The social and instrumental support components involves training mothers in decision-making and problem-solving skills, for example maintaining adequate food with limited financial resources. **Delivered by** – The intervention team consisted of		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	a therapist and one or more advanced undergraduate or postbaccalaureate students (11 masters-level licensed mental health service providers were hired and trained, supervised by a clinical psychologist). **Delivered to** – The intervention was delivered to mothers who were randomly assigned to the Project Support condition (n=17). **Duration, frequency, intensity** – The intervention was designed to include weekly sessions of 1:1 1/2hr for up to 8 months. There was not a specific set number of sessions - the intervention was structured so that it could be delivered flexibly within the 8 month period. **Key components and objectives of intervention** – The intervention sought to improve parenting skills and reduce psychological distress. **Content/session titles** – Therapists addressed mothers' beliefs, practices, and knowledge about parenting, in addition to their children's behaviour patterns. Included among the 12 skills that were taught		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	were: attentive and nondirective play with your child, and listening to and comforting your child. Therapists monitor mothers' mastery of the parenting skills – any skills not fully mastered are revisited to ensure mastery of each skill before moving on to the next. The social and instrumental support component of the intervention included training in decision-making and problem-solving skills. **Location/place of delivery** – Sessions were delivered at families' homes. **Describe comparison intervention** – Services as usual varied considerably across the 18 families. Four did not receive any services and of the 14 who did, all received some type of parenting intervention. Twelve families also received services in addition to parenting, including anger management, GED classes and individual therapy. **Outcomes measured** Incidence of abuse and neglect – CPS records were reviewed to assess whether participating families had again been referred to CPS for child maltreatment during the 20		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	months following their baseline assessment. Quality of parenting and parent-child relationships – Mothers' perceived inability to manage childbearing responsibilities was measured using the Parenting Locus of Control Scale (PLOC; Campis et al. 1986). Mothers' reports of harsh parenting behaviours were measured using the psychological aggression and minor assault subscales from the Revised Conflict Tactics Scales (CTS-R; Straus et al. 1996). Ineffective parenting was also observed by staff. Caregiver/parent health and wellbeing outcomes – Maternal psychological distress was measured using the Symptom Checklist-90-Revised (SCL-90-R; Derogatis et al. 1976).		

10a. Lieberman AF, Van Horn PJ, Ghosh Ippen C (2005) Toward evidence-based treatment: Child-parent psychotherapy with preschoolers exposed to marital violence. Journal of the American Academy of Child and Adolescent Psychiatry 44: 1241–8

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim – The trial sought to evaluate the effectiveness of Child Parent	**Participants** **Children and young people** – Participants included 39 girls and 36 boys aged 3-5 (M=4.06, SD=.82) who	**Effect sizes** **Children and young people's health and wellbeing outcomes**	**Overall assessment of internal validity** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Psychotherapy (CPP) in preschoolers who had been exposed to 'marital violence'. The results are reported in three papers. One paper reports results up to conclusion of treatment (Lieberman et al. 2005), one at 6 months post-treatment (Lieberman et al. 2006) and 1 considers effectiveness in relation to exposure to multiple traumatic and stressful events (TSEs) (Ghosh et al. 2011). **Methodology** RCT. **Country** Not UK. – US study.	had been exposed to marital violence, as confirmed by mother's report. (NB under the Adoption and Children Act 2002, significant harm includes 'impairment suffered from seeing or hearing the ill-treatment of another'.) **Caregivers and families** – Mothers of the 39 girls and 36 boys were also included (the exact number is not stated), forming child-parent dyads. **Sample characteristics** Age – Children were aged between 3 and 5 (M=4.06, SD=.82). Mothers averaged 31.48 years (SD=6.23). Sex – There were 39 girls, 36 boys and their mothers. Ethnicity – 38.7% of children were mixed ethnicity (predominantly Latino/White), 28% Latino, 14.7% African American, 9.3% White, 6.7% Asian and 2.6% other. Mothers were 37.3% Latina, 24% White, 14.7% African American, 10.7% Asian and the rest were mixed race or other ethnicities. Religion/belief – Not stated. Disability – Children and mothers with 'mental	– Child functioning - semistructured interview for diagnostic classification DC (Scheeringa et al. 1995) There was a significant group x time interaction regarding the total number of traumatic stress disorder symptoms, in favour of the intervention, with medium effect size - F(1,59)=10.98, p<0.001, d=0.63. Follow-up analyses showed that the CPP group had a significant reduction in traumatic stress disorder symptoms - t(32)=5.46, p<0.001), whereas the comparison group did not. Child Behavior Checklist (CBCL) There was a significant group x time interaction for CBCL scores in favour of the intervention, with small effect size - F(1,61)=5.77, p<0.05, d=0.24. Follow-up analyses showed that only the CPP group showed significant reductions in CBCL scores (t(34)=2.86, p<0.01. Clinical significance of the effects was examined by analysis of the percentage of children in each group meeting the criteria for a diagnosis of Traumatic Stress Disorder. At intake, the two groups were not significantly different. At post-test, there was a statistically significant group difference in favour of CPP (chi-square (n=61) = 8.43, p<0.01, phi=0.37). The authors state that intent-to-treat analyses resulted in similar results to the above, but these results are not reported. **Caregiver/parent health and wellbeing outcomes** – The Symptoms Checklist-90 Revised (Derogatis, 1994) Global Severity Index: There was a marginally significant time x group interaction in favour of CPP, with a small to medium effect size - F(1, 59)=3.48, p=0.07, d=0.037). Follow-up analyses showed that the CPP group showed statistically significant reduction in GSI (t(32)=4.47, p<0.001), whereas for comparison	**Overall assessment of external validity** + **Overall validity score** + Small Sample size - Reliance on maternal report - Short follow-up period - In Ghosh (2011), query validity of dichotomisation of children into <4 and 4+ TSE risk groups (as nearly all children in the <4 group had experienced at least 2 TSEs).

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	retardation' (Lieberman et al. 2005, p1243) were excluded from the study, as were children with autistic spectrum disorder. Long term health condition – Not reported. Sexual orientation – Not reported. Socioeconomic position – Mean monthly family income was $1,817 (SD=$1,460; range $417–$8,333). Public assistance was received by 23% of families and 41% had incomes below the federal poverty level, according to the Department of Health and Human Services Guidelines (2004). Type of abuse – Child–mother dyads were recruited if the child had been exposed to marital violence, as confirmed by mother's report on the Conflict Tactics Scale 2 (Straus et al. 1996). The prevalence of exposure to the 8 TSEs was as follows: physical abuse (29.3%); sexual abuse (12%); witnessing domestic violence (97.3%); neglect (5%); separation from a caregiver (100%); caregiver criminal history (5.3%); caregiver substance abuse (16%); and caregiver mental illness (88%). High prevalence rates of domestic violence exposure and separation from a caregiver are related to study criteria. In Ghosh et al.	group this was only marginally significant (t(27)=1.94, p=0.06). Clinician-administered PTSD Scale (CAPS) CAPS re-experiencing - time x group interaction was not significant, showing that intervention was not significantly better than control (d=0.29). CAPS avoidance - there was a significant time x group interaction in favour of the intervention, with medium effect size - F(1, 57)=5.08, p<0.05, d=0.50. Follow-up analyses showed significant intake-outcome reductions in avoidance symptoms for CPP group only (t(33=5.16, p<0.001). CAPS hyperarousal - time x group interaction was not significant, showing that intervention was not significantly better than control (d=0.19). CAPS total - time x group interaction was marginally significant, with small to medium effect size - F(1, 57)=3.23, p<0.1, d=0.41. Follow-up analyses showed that both CPP and comparison groups showed significant intake-outcome reductions (CPP: t(33)=5.34, p<0.001; Comparison: t(24)=2.50, p<0.05). Clinical significance of treatment effects was explored by examining percentage of mothers in each group with a diagnosis of PTSD. At intake there was no group difference, at outcome there was also no statistically significant difference in PTSD rates across the two groups (chi-square (n=60)=2.26, p=non-significant, phi=0.19). The authors state that intent-to-treat analyses resulted in similar results to the above, but these results are not reported. **Narrative findings** Narrative findings - effectiveness – Child Participants in the CPP group showed a significantly greater reduction in the number of traumatic	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	(2011) the number of traumatic and stressful events to which children had been exposed was also recorded. Of the children, 12% had experienced 2 TSEs, 41.3% had 3 TSEs, and 46.7% had 4+ TSEs. Looked after or adopted status – Not reported. Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size** Comparison numbers – n=33, (in Ghosh et al. 2011, 15 of whom had been exposed to >4 TSEs, and 18 to <4 TSEs). Intervention number – n=42, (in Ghosh et al. 2011, 20 of whom had been exposed to >4 TSEs and 22 to <4 TSEs). **Sample size** – A total of 75 child-mother dyads were included in the study. **Intervention category** Child-Parent Psychotherapy. **Intervention** **Describe intervention** – CPP is an empirically supported treatment, based on the premise that	stress disorder symptoms than those in the intervention group. They also showed a greater reduction in behaviour problems, as measured by the Child Behavior Checklist, and were significantly less likely to be diagnosed with Traumatic Stress Disorder. There was a trend towards mothers in the CPP group showing a significantly greater improvement in maternal functioning, as measured by the Symptoms Checklist-90 Revised (Derogatis 1994) Global Severity Index, although this was not statistically significant. There was also a trend towards mothers in the CPP groups showing a significantly greater improvement in PTSD symptoms, as measured by the clinician-administered PTSD Scale, although this was not statistically significant, but did have a small to medium effect size. No significant differences were seen on the re-experiencing, and hyperarousal subscales. There was a significant difference in favour of the intervention on the avoidance subscale. There was no difference in rates of PTSD diagnoses at post-treatment between the CPP and comparison groups.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	the attachment system is the main organiser of children's responses to danger and safety in the first years of life. CPP targets maladaptive behaviours, supports developmentally appropriate interactions, and guides both child and mother in creating a joint narrative of the traumatic events while working toward their resolution. **Delivered by** – CPP was delivered by clinicians who had at least a masters degree in clinical psychology. **Delivered to** – The intervention was delivered to dyads randomly assigned to the CPP condition (n= 22). **Duration, frequency, intensity** – Weekly CPP sessions lasted approximately 60 minutes and were conducted over the course of 50 weeks. Dyads attended a mean of 32.09 sessions (SD=15.20). **Key components and objectives of intervention** – CPP is guided by the unfolding child–parent interactions, with the therapeutic goal of enhancing parents' ca-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	pacity to provide safety and developmentally appropriate caregiving to the child. **Content/session titles** – The treatment manual included clinical strategies and clinical illustrations to address the following domains of functioning: play; sensorimotor disorganization and disruption of biological rhythms; fearfulness; reckless, selfendangering, and accident-prone behaviour; aggression; punitive and critical parenting; and the relationship with the perpetrator of the violence and/or absent father (taken from original study; Lieberman et al. 2005). **Location/place of delivery** – Not reported. **Describe comparison intervention** – Comparison mothers received individual psychotherapy plus case management from a PhD degree-level clinician. This involved at least monthly phone calls and inquiries about how mother and child were doing. Face-to-face meetings were also scheduled when clinically indicated. **Outcomes measured** Children and young people's health and wellbeing outcomes		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	– Child functioning - the semistructured interview for diagnostic classification DC (Scheeringa et al. 1995) was used to measure the number of PTSD and depression symptoms the child was experiencing. The Child Behavior Checklist (CBCL; Achenbach 1991, 1992) was additionally completed with parents complete to detect emotional and behavioural problems in children. Caregiver/parent health and wellbeing outcomes – The Symptoms Checklist-90 Revised (Derogatis 1994) was used to assess maternal functioning Clinician-administered PTSD Scale (CAPS) was used to measure maternal PTSD.		

10b. Lieberman AF, Ghosh Ippen C, Van Horn P (2006) Child-parent psychotherapy: 6-month follow-up of a randomized controlled trial. Journal of the American Academy of Child and Adolescent Psychiatry 45: 913–18

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim The trial sought to evaluate the effectiveness of Child Parent Psychotherapy (CPP) in preschool children who had been exposed to 'marital violence'. The results	**Participants** **Children and young people** – Children between the ages of 3 and 5 who had been exposed to marital violence, as confirmed by mother's report. (NB under the Adoption and Children Act 2002, significant harm in-	**Effect sizes** **Children and young people's health and wellbeing outcomes** – Total Behavior Problem scale of the Child Behavior Checklist - Treatment Completer analyses - Significant group x time interaction with a small to medium effect size, in favour of the intervention group – F (1, 48)=5.39, p < .05, d=0.41. Significant main effect for	**Overall assessment of internal validity** + **Overall assessment of external validity** ++ **Overall validity score** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
are reported in three papers. One paper reports results up to conclusion of treatment (Lieberman et al. 2005), one at 6 months post-treatment (Lieberman et al. 2006) and 1 considers effectiveness in relation to exposure to multiple traumatic and stressful events (TSEs) (Ghosh et al. 2011). **Country** Not UK, USA. **Source of funding** Government – National Institute of Mental Health. Voluntary/charity – Irving Harris Foundation.	cludes 'impairment suffered from seeing or hearing the ill-treatment of another'.) **Caregivers and families** – Mothers of children between the ages of three and five who had been exposed to marital violence, forming child-parent dyads. **Sample characteristics** Age – Children – Originally randomised – 3-5 years old (M=4.06, SD=.82). Final six month follow-up sample – 3-6 years old (M=4.04, SD=.82). Age of mothers not reported. Sex – 39 girls and 36 boys (and their mothers) were originally randomised. Final sample 6 month follow-up sample included 22 girls and 28 boys. Ethnicity – Children – Originally randomised – 38.7% of children were mixed ethnicity (predominantly Latino/White), 28% Latino, 14.7% African American, 9.3% White, 6.7% Asian and 2.6% other. Mothers were 37.3% Latina, 24% White, 14.7% African American, 10.7% Asian and the rest were mixed race or other ethnicities. Final sample 6 month	time - F (1, 48)=14.35, p<.001. Follow-up analyses showed that only the intervention group evidenced significant reductions – t (26)=3.92, p< 001. Total Behavior Problem scale of the Child Behavior Checklist - Intent to treat analyses – Significant group x time interaction with a small to medium effect size, in favour of the intervention group - F (1, 73)=5.44, p<.05, d=0.44. Significant main effect for time - F (1, 73)=14.08, p<.001. Follow-up analyses showed that only the intervention group evidenced significant reductions – t (41)=4.07, p<.001. Caregiver/parent health and wellbeing outcomes – Global Severity Index of the Symptoms Checklist-90 Revised - Treatment Completer analyses - Significant group x time interaction with a small to medium effect size, in favour of the intervention group – F (1, 47)=5.12, p=.05, d=0.38. Significant main effect for time - F (1, 47)=21.50, p<.001. Follow-up analyses showed that only the intervention group evidenced significant reductions – t (26)=5.11, p<.001. Global Severity Index of the Symptoms Checklist-90 Revised - Intent to treat analyses – Group x time interaction – Not reported. Significant main effects for time - F (1, 73)=14.92, p<.001. **Narrative findings - effectiveness** – Children and young people's health and wellbeing outcomes – Treatment completer analysis - At 6 months post-intervention, children in the intervention group showed a significantly greater reduction in behavioural problems than those in the comparison group as measured by the Total Behavior Problem	The failure to present intent-to-treat analysis of maternal outcomes and use of measures which rely on maternal report are important limitations of the study.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	follow-up sample – 38% mixed ethnicity, 28% Latino, 16% African American, 12% white, 4% Asian, and 2% of another ethnicity. Religion/belief – Not reported. Disability – Children and mothers with 'mental retardation' (Lieberman et al. 2005: 1243) were excluded from the study, as were children with autistic spectrum disorder. Long term health condition – Not reported. Sexual orientation – Not reported. Socioeconomic position – Not reported (further demographic details provided in Lieberman et al. 2005). Type of abuse – Child–mother dyads were recruited if the child had been exposed to marital violence, as confirmed by mother's report on the Conflict Tactics Scale 2 (Straus et al. 1996). Further details on exposure to traumatic and stressful events are provided in Lieberman et al. 2005). Looked after or adopted status – Not reported.	scale of the Child Behavior Checklist. Follow-up analyses showed that only children assigned to the intervention group showed significant reductions in scores on this measure. Intent to treat analysis. At 6 months post-intervention, children in the intervention group showed a significantly greater reduction in behavioural problems than those in the comparison group as measured by the Total Behavior Problem scale of the Child Behavior Checklist. Follow-up analyses showed that only children assigned to the intervention group showed significant reductions in scores on this measure. Caregiver/parent health and wellbeing outcomes – Treatment completer analysis - At 6 months post-intervention, mothers in the intervention group showed a significantly greater reduction in severity of psychiatric symptoms (functioning) than those in the comparison group as measured by the Global Severity Index of the Symptoms Checklist-90 Revised. Follow-up analyses showed that only mothers assigned to the intervention group showed significant reductions in scores on this measure. Intent to treat analysis – Not reported.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size** Comparison numbers – Originally randomised n=33, (in Ghosh et al. 2011, 15 of whom had been exposed to >4 TSEs, and 18 to <4 TSEs). Final six month follow-up sample n=23. Intervention number – Originally randomised n=42, (in Ghosh et al. 2011, 20 of whom had been exposed to >4 TSEs and 22 to <4 TSEs.) Final six month follow-up sample n=27. Sample size – Originally randomised n=75. Final six month follow-up sample n=50 (dyads who completed treatment and were included in Treatment Completer analyses). **Intervention category** Child-Parent Psychotherapy. **Intervention** **Describe intervention** – Child-Parent Psychotherapy is an empirically supported treatment, based on the premise that the attachment		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	system is the main organiser of children's responses to danger and safety in the first years of life. The intervention targets maladaptive behaviours, supports developmentally appropriate interactions, and guides both child and mother in creating a joint narrative of the traumatic events while working toward their resolution. **Delivered by** – Child-Parent Psychotherapy was delivered by clinicians who had at least a Master's degree in clinical psychology. **Delivered to** – The intervention was delivered to parent-child dyads randomly assigned to the CPP condition (n=42 were originally randomised, final 6 month follow-up sample n=27). **Duration, frequency, intensity** – Weekly Child-Parent Psychotherapy sessions lasted approximately 60 minutes and were conducted over the course of 50 weeks. Dyads attended a mean of 32.09 sessions (SD=15.20). **Key components and objectives of intervention** – Child-Parent Psychotherapy is guided by the unfolding child–parent		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	interactions, with the therapeutic goal of enhancing parents' capacity to provide safety and developmentally appropriate caregiving to the child. **Content/session titles** – The treatment manual included clinical strategies and clinical illustrations to address the following domains of functioning: play; sensorimotor disorganization and disruption of biological rhythms; fearfulness; reckless, self-endangering, and accident-prone behaviour; aggression; punitive and critical parenting; and the relationship with the perpetrator of the violence and/or absent father (taken from original study; Lieberman et al. 2005). **Location/place of delivery** – Not reported. **Describe comparison intervention** – Comparison mothers received individual psychotherapy plus case management from a PhD degree-level clinician. This involved at least monthly phone calls and inquiries about how mother and child were doing. Face-to-face meetings were also scheduled when clinically indicated. The authors report that in '… the comparison group, 73% (n=22) of mothers and		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	55% (n=17) of children received individual treatment, and 45% (n=14) received separate individual psychotherapy for both mother and child. Mothers in the comparison group reported a range of 2 to 50 sessions for children and 6 to 50 sessions for themselves, with 50% of the mothers and 65% of the children receiving more than 20 individual sessions' (p915). **Outcomes measured** Children and young people's health and wellbeing outcomes – Stress-related behaviours (parental reports of emotional and behavioural problems) were measured using the Total Behavior Problem scale of the Child Behavior Checklist (Achenbach 1991, 1992). Caregiver/parent health and wellbeing outcomes – Caregiver/parent health and wellbeing outcomes The Global Severity Index of the Symptoms Checklist-90 Revised (Derogatis, 1994) was used to assess maternal functioning (current psychiatric symptoms).		

11c. Ghosh Ippen C, Harris WW, Van Horn P et al. (2011) Traumatic and stressful events in early childhood: Can treatment help those at highest risk? Child Abuse and Neglect 35: 504–13

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim – This study is a reanalysis of data from the Lieberman et al. (2005, 2006) treatment outcome and follow-up studies, which examined the effectiveness of Child Parent Psychotherapy (CPP) in preschoolers who had been exposed to multiple traumatic and stressful events (TSEs). The aim of this study was to examine the effects of CPP in children with 4+ TSEs, and whether level of child risk influences treatment effects on maternal symptoms. **Methodology** RCT – Participants were randomly allocated to one of two conditions: the intervention (CPP) or individual psychotherapy plus case management.	**Participants** **Children and young people** – Participants included 39 girls and 36 boys aged 3-5 (M=4.06, SD=.82) who had been exposed to marital violence, as confirmed by mother's report. (NB under the Adoption and Children Act 2002, significant harm includes 'impairment suffered from seeing or hearing the ill-treatment of another'.) **Caregivers and families** – Mothers of the 39 girls and 36 boys were also included (the exact number is not stated), forming child-parent dyads. **Sample characteristics** Age – Children were aged between 3 and 5 (M=4.06, SD=.82). Mothers averaged 31.48 years (SD=6.23). Sex – There were 39 girls, 36 boys and their mothers. Ethnicity – 38.7% of children were mixed ethnicity (predominantly Latino/White), 28% Latino, 14.7% African American, 9.3% White, 6.7% Asian and 2.6% other. Mothers were 37.3% Latina, 24%	**Effect sizes** **Children and young people's health and wellbeing outcomes** – All analyses were conducted for both the full sample (intent to treat analysis, ITT) and treatment completers only (TC). We have extracted data for the ITT calculations only, as these provide a more conservative analysis. Main effects of time and interactions of time and TSE status are not reported here, as they do not provide information regarding treatment effectiveness. PTSD – pre to post-test. There was a significant time x treatment interaction effect (F= 14.71, p<0.001, eta-squared=0.17) in favour of the intervention group. There was also a significant time x treatment x TSE interaction effect (F=3.99, p<0.05, eta-squared=0.05) in favour of the intervention group, with greater improvements observed for children with 4+ TSEs in the CPP group. To examine clinically significant reductions in PTSD symptoms, chi-square tests were used to compare the number of CPP and comparison group children diagnosed with PTSD. At posttest, there were statistically significant group differences for children with 4+ TSEs, with children in the CPP group showing significantly lower rates of PTSD (chi-square(1)=10.48, p<0.01, phi=0.55). Depression – pre to posttest There was a significant time x treatment interaction effect (F=4.34, p<0.05, eta-squared = 0.06) in favour of the intervention group. There was also a significant time x treatment x TSE interaction effect (F=4.52, p<0.05, eta-squared = 0.06) in	**Overall assessment of internal validity** + **Overall assessment of external validity** + **Overall validity score** + Key limitations of the study include: - Small sample size - Reliance on maternal report - Short follow-up period - No reported blinding of assessors - Dichotomisation of children into <4 and 4+ TSE risk groups (as nearly all children in the <4 group had experienced at least 2 TSEs).

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Country Not UK. – US study. **Source of funding** Other – This research was supported by the National Institute of Mental Health Grant and the Coydog Foundation.	White, 14.7% African American, 10.7% Asian and the rest were mixed race or other ethnicities. Religion/belief – This is not stated. Disability – Children and mothers with 'mental retardation' (p506) were excluded from the study, as were children with autistic spectrum disorder. Long term health condition – This is not reported. Sexual orientation – This is not reported. Socioeconomic position – Mean monthly family income was $1,817 (SD=$1,460; range $417–$8,333). Public assistance was received by 23% of families and 41%had incomes below the federal poverty level, according to the Department of Health and Human Services Guidelines (2004). Type of abuse – Child–mother dyads were recruited if the child had been exposed to marital violence, as confirmed by mother's report on the Conflict Tactics Scale 2 (Straus et al., 1996). The prevalence of exposure to the 8 TSEs was as follows: physical abuse (29.3%); sexual abuse (12%); witnessing domestic vio-	favour of the intervention group, with greater improvements observed for children with 4+ TSEs in the CPP group. Co-occurring diagnoses – pre to post-test. There was a marginally significant time x treatment interaction effect (F=2.86, p<0.1, eta-squared =0.04) in favour of the intervention group. There was no significant time x treatment x TSE interaction. Total CBCL – pre to posttest There was a significant time x treatment interaction effect (F=7.25, p<0.01, eta-squared = 0.09) in favour of the intervention group. There was also a significant time x treatment x TSE interaction effect (F=6.83, p<0.05, eta-squared = 0.09), with greater improvements observed for children with 4+ TSEs in the CPP group. Total CBCL – pre to 6 month follow up There was a significant time x treatment interaction effect (F=7.47, p<0.01, eta-squared = 0.10) in favour of the intervention group. There was also a significant time x treatment x TSE interaction effect (F=12.19, p<0.001, eta-squared =0.15), with greater improvements observed for children with 4+ TSEs in the CPP group. **Caregiver/parent health and wellbeing outcomes** – PTSD – pre to post-test. There was a significant time x treatment group interaction effect (F=3.98, p<0.05, eta-squared = 0.05) in favour of the intervention. There was no significant time x treatment x TSE interaction effect. Post hoc tests found that mothers in the CPP group showed a significant improvement from pre to posttest, for mothers of both <4 TSE children (t(21)=3.81, p<0.01, d=0.68) and 4+ children (t(21) =3.17, p<0.01, d=0.92), whereas in the comparison	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	lence (97.3%); neglect (5%); separation from a caregiver (100%); caregiver criminal history (5.3%); caregiver substance abuse (16%); and caregiver mental illness (88%). High prevalence rates of domestic violence exposure and separation from a caregiver are related to study criteria. Of the children, 12% had experienced 2 TSEs, 41.3% had 3 TSEs, and 46.7% had 4+ TSEs. Looked after or adopted status – This is not stated. Unaccompanied asylum seeking, refugee or trafficked children – This is not stated. **Sample size** Comparison numbers – n=33, 15 of whom had been exposed to >4 TSEs, and 18 to <4 TSEs. Intervention number – n=42, 20 of whom had been exposed to >4 TSEs and 22 to <4 TSEs. Sample size – A total of 75 child-mother dyads were included in the study. **Intervention** **Describe intervention** – CPP is an empirically supported treatment, based on the premise that	group only mothers of <4 children made a significant improvement (t(17)=2.55, p<0.05, d=0.76). Chi-squared analyses found that, at posttest, in the ITT sample, CPP 4+ mothers were significantly less likely to have a diagnosis of PTSD [X2(1)=7.70, p=.01., phi = .47], with 15% of CPP mothers and 60% of comparison group mothers meeting PTSD criteria. No significant treatment differences for maternal PTSD were found for the <4 group. Depression - pre to posttest There was a marginally significant time x treatment group interaction effect ((F=3.76, p<0.1, eta-squared = 0.05). There was no significant time x treatment x TSE interaction effect. **Narrative findings - effectiveness** – Overall, children in the CPP group showed significantly greater improvements on PTSD, depression, co-occurring diagnoses and behaviour problems, compared to those in the comparison group. The rate of improvement was greater amongst children with 4+ TSEs for all these variables. CPP also had a greater impact on PTSD and depression in mothers, compared to the comparison group. There is some evidence that the beneficial effects on PTSD are more pronounced for mothers of higher risk (4+ TSE children). The authors note that 'In the <4 TSE group, both treatment and comparison group mothers showed significant improvements in PTSD, but only CPP mothers showed significant posttreatment reductions in depression…In the 4+ TSE group, CPP mothers showed significant reduction in PTSD and depression … whereas comparison group mothers showed no improvements in any of these domains' (p510).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	the attachment system is the main organiser of children's responses to danger and safety in the first years of life. CPP targets maladaptive behaviors, supports developmentally appropriate interactions, and guides both child and mother in creating a joint narrative of the traumatic events while working toward their resolution. **Delivered by** – CPP was delivered by clinicians who had at least a masters degree in clinical psychology. **Delivered to** – The intervention was delivered to dyads randomly assigned to the CPP condition (n= 22). **Duration, frequency, intensity** – Weekly CPP sessions lasted approximately 60 minutes and were conducted over the course of 50 weeks. Dyads attended a mean of 32.09 sessions (SD= 15.20). **Key components and objectives of intervention** – CPP is guided by the unfolding child–parent interactions, with the therapeutic goal of enhancing parents' ca-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	pacity to provide safety and developmentally appropriate caregiving to the child. **Content/session titles** – The treatment manual included clinical strategies and clinical illustrations to address the following domains of functioning: play; sensorimotor disorganization and disruption of biological rhythms; fearfulness; reckless, self-endangering, and accident-prone behaviour; aggression; punitive and critical parenting; and the relationship with the perpetrator of the violence and/or absent father (taken from original study; Lieberman et al. 2005). **Location/place of delivery** – This is not stated. **Describe comparison intervention** – Comparison mothers received individual psychotherapy plus case management from a PhD degree-level clinician. This involved at least monthly phone calls and inquiries about how mother and child were doing. Face-to-face meetings were also scheduled when clinically indicated. **Outcomes measured** Children and young people's health		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	and wellbeing outcomes – The semistructured interview for diagnostic classification DC (Scheeringa et al. 1995) was used to measure the number of PTSD and depression symptoms the child was experiencing. The Child Behavior Checklist (CBCL; Achenbach 1991, 1992) was additionally completed with parents complete to detect emotional and behavioural problems in children. Caregiver/parent health and wellbeing outcomes – The Symptoms Checklist-90 Revised (Derogatis 1994) was used to assess maternal functioning and the Clinician-administered PTSD Scale (CAPS) was used to measure maternal PTSD.		

12. Linares LO, Li MM, Shrout PE (2012) Child training for physical aggression? Lessons from foster care. Children and Youth Services Review 34: 2416–22

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim – The study aimed to evaluate the effectiveness of an adapted version of the Incredible Years Child Training programme in reducing	**Participants** **Children and young people** – Foster children between the ages of 5 and 8 with substantiated neglect. At baseline 49% met criteria for an externalising disorder and 39% met criteria	**Effect sizes** Children and young people's health and wellbeing outcomes. – Primary multilevel analysis (post-intervention to follow-up). Effect sizes not reported in the paper and have not been calculated by review team due to differences in the 2 groups at baseline.	**Overall assessment of internal validity** + **Overall assessment of external validity** ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
physical aggression in young children in foster homes. **Country** Not UK. – USA - New York. **Source of funding** Not reported.	for Attention Deficit Hyperactivity Disorder. There do not appear to have been any eligibility criteria in relation to physical aggression. **Caregivers and families** – Biological and foster parents of foster children between the ages of 5 and 8. **Professionals/practitioners** – Teachers of foster children between the ages of 5 and 8. **Sample characteristics** Age – Mean age of child at baseline assessment (standard deviation) - Intervention 6.7 years (1.1), control 6.7 years (1.3). Mean age of child at placement in foster care (standard deviation) - Intervention 4.9 years (1.8), control 4.7 years (2.4). Sex – Male children - Intervention n=29 (59%), control n=17 (38%). Ethnicity – African American - intervention n=18 (37%), control n=28 (62%); Latino - intervention n=14 (29%), control n=9 (20%), 'Other (mixed, Caucasian, other)' intervention n=17 (35%), control n=8 (18%).	Foster parent ratings of physical aggression measured using 6 items from the child behavior checklist aggression subscale: Main effect of time – Significant effect for time – foster parent reports of physical aggression in both groups declined over time – estimate = -1.47, p<.01. Group x time interaction effect (improvement from post-intervention to follow-up) – Significant effect in favour of the control group - estimate = -1.41, p<.01. Foster parent ratings of good self-control: Main effect of time – Not reported. Treatment group – Significant difference in favour of the control group - estimate = -.27, p<.05. Group x time interaction effect (improvement from post-intervention to follow-up) - Significant effect in favour of the control group - estimate = -.33, p<.05. Foster parent ratings of poor self-control No significant differences from post-intervention to follow up Classroom teacher ratings Statistical analyses of teacher ratings are not reported as these showed no change. Moderator analyses Foster parent ratings of physical aggression measured using 6 items from the child behavior checklist aggression subscale (post-intervention – controlling for baseline scores, child ethnicity, adhd diagnosis and study site): Group x gender: Significant effect for group x gender with male children in the intervention group showing lower scores than those in the control group. Foster parent ratings of self-control (post-intervention): Group x type of maltreatment interaction: Significant effect for group x type of maltreatment interaction with children in the control group with a history of neglect (in the presences of abuse) showing higher levels of self-control than those in the intervention group.	***Overall validity score*** + – A limitation of the study is the short follow-up period (3 months).

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Religion/belief – Not reported. Disability – Not reported. Long term health condition – Not reported. Sexual orientation – Not reported. Socioeconomic position – Not reported. Type of abuse – Overall maltreatment was identified in the case records of 97% of the sample. Type of maltreatment – Physical abuse - Intervention n=11 (22%), control n=9 (20%). Sexual abuse - Intervention n=0 (0%), control n=3 (7%). Type of neglect – Lack of supervision - Intervention n=30 (61%), control n=31 (69%). Failure to provide - Intervention n=21 (43%), control n=17 (38%). Failure to protect (Domestic violence) - Intervention n=14 (29%), control n=10 (22%). Emotional - Intervention n=13 (27%), control n=11 (24%). Educational - Intervention n=8 (16%), control n=7 (16%). Legal/moral - Intervention n=3 (6%), control n=4 (9%). Any neglect - Intervention n=46 (94%), control n=38 (80%). Looked after or adopted status – All children were living in foster homes. Placement in a kinship foster	**Narrative findings** Narrative findings - effectiveness – Foster parent ratings of physical aggression measured using 6 items from the child behavior checklist aggression subscale: The study found a group x time interaction effect (improvement from post-intervention to follow-up) in favour of the control group, meaning that those in Usual Care showed more improvement from post-intervention to follow-up than those in the Child Training group. Foster parent ratings of self-control: There was a significant main effect of treatment group in favour of the control group, with foster parents of the Usual Care group rating the child's self-control higher than those in the intervention group at both post-intervention and follow-up. There was significant group x time interaction effect (improvement from post-intervention to follow-up) in favour of the control group, with foster parent reports of the child's self-control showing more improvement for those in Usual Care compared to the Child Training group. The study examined two variables which may interact with treatment condition. It was found that there was a significant interaction with gender: boys in the Child Training group at post intervention showed lower ratings of physical aggression that boys in the Usual Care group. There was also a significant interaction with type of maltreatment: children in the Usual Care group with histories of neglect showed higher good elf control than those in the Child Training group. The study reports that teacher ratings across the three outcomes 'remain unchanged'. No statistical data reported. **Effect sizes** **Children and young people's health and wellbeing**	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	home - Intervention n=18 (37%), control n=19 (42%). Mean number of children in the foster home (standard deviation) – Intervention 2.6 (1.5), control 2.9 (1.2). Mean number of weekly family visits (standard deviation) – Intervention 1.1 (2.0), control 1.1 (1.0). Cases in which parental rights had been terminated – Intervention n=5 (10%), control n=5 (11%). Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size** Comparison numbers – n=45 children. Intervention number – n=49 children. **Sample size** – n=94 eligible children were randomised. Intervention - n=49, control n=45. **Intervention category** Other. – Child training programme - A '… a child-focused adaptation of the Incredible Years Child Training program …' (p2416). **Intervention** **Describe intervention**	**outcomes** Primary multilevel analysis (post-intervention to follow-up) Foster parent ratings of physical aggression measured using 6 items from the child behavior checklist aggression subscale: Main effect of time – Significant effect for time – foster parent reports of physical aggression in both groups declined over time – estimate = -1.47, p<.01. Group x time interaction effect (improvement from post-intervention to follow-up) – Significant effect in favour of the control group - estimate = -1.41, p<.01. Foster parent ratings of good self-control: Main effect of time – Not reported. Treatment group – Significant difference in favour of the control group - estimate = -.27, p<.05. Group x time interaction effect (improvement from post-intervention to follow-up) - Significant effect in favour of the control group - estimate = -.33, p<.05. Foster parent ratings of poor self-control: No significant differences from post-intervention to follow up Classroom teacher ratings - Statistical analyses of teacher ratings are not reported as these showed no change. Moderator analyses Foster parent ratings of physical aggression measured using six items from the child behavior checklist aggression sub scale (post-intervention – controlling for baseline scores, child ethnicity, ADHD diagnosis and study site): Group x gender: Significant effect for group x gender with male children in the intervention group showing lower scores than those in the control group. Foster parent ratings of self-control (post-intervention): Group x type of maltreatment interaction: Significant	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	– 'Treatment' version of the Incredible Years Training programme (manualised). The authors report that they selected 12 out of a possible 18 lessons from the Incredible Years Dina Program for Young Children which they believed would enable them to target the self-regulation processes of the participating children whilst remaining feasible. The goals of this intervention are not clearly specified but the authors note that they selected the treatment version because it had previously been found to be effective in reducing conduct problems and conflicts with peers for 4 to 8 year old children with conduct disorders, Attention Deficit Hyperactivity disorder and Oppositional Defiant disorder, noting the high prevalence of these types of disorders in foster children. Delivered by – A team of 3 clinicians (one from a university and 2 from the agency through which the child was accessing services) with at least a masters level qualification in psychology or social work. Delivered to – Foster children between the ages of 5 and 8 in groups of 6 to 9. Foster parents (and biological parents where available) also participated in group	effect for group x type of maltreatment interaction with children in the control group with a history of neglect (in the presences of abuse) showing higher levels of self-control than those in the intervention group. **Narrative findings - effectiveness** – Foster parent ratings of physical aggression measured using 6 items from the child behavior checklist aggression sub scale: The study found a group x time interaction effect (improvement from post-intervention to follow-up) in favour of the control group, meaning that those in Usual Care showed more improvement from post-intervention to follow-up than those in the Child Training group. Foster parent ratings of self-control: There was a significant main effect of treatment group in favour of the control group, with foster parents of the Usual Care group rating the child's self-control higher than those in the intervention group at both post-intervention and follow-up. There was significant group x time interaction effect (improvement from post-intervention to follow-up) in favour of the control group, with foster parent reports of the child's self-control showing more improvement for those in Usual Care compared to the Child Training group. The study examined 2 variables which may interact with treatment condition. It was found that there was a significant interaction with gender: boys in the Child Training group at post intervention showed lower ratings of physical aggression than boys in the Usual Care group. There was also a significant interaction with type of maltreatment: children in the Usual Care	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	classes at sessions 1, 6 and 12. **Duration, frequency, intensity** – The programme delivered to children consisted of 12 2-hour sessions; parent sessions also lasted for 2 hours. (No further details are provided). **Key components and objectives of intervention** – Little detail on the intervention is provided, only that it addresses the self-regulatory processes of the child. The authors note that they added a lesson to the programme which aimed to enhance the child's sense of 'belongingness' to their foster home. They also report that parent sessions were '… aimed at promoting skill generalisation to the foster home (or during the family visitation) and assist in homework activities' (p2418). **Content/session titles** – The authors report that the Incredible Years modules which they used were 'Understanding and Detecting Feelings', 'Detective Wally Teaches Problem Solving Steps', and 'Tiny Turtle Teaches Anger Management'. The 'My Homes, My Families' lesson was a lesson developed by the authors and aimed to enhance the child's sense of 'belongingness' to their foster home.	group with histories of neglect showed higher good elf control than those in the Child Training group. The study reports that teacher ratings across the 3 outcomes 'remain unchanged'. No statistical data reported.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Location/place of delivery – Classroom like settings at 1 of 6 '… volunteering community sites …' (p2417). **Describe comparison intervention** – Care as usual - no information regarding the services which the comparison group received are provided. **Outcomes measured** Children and young people's health and wellbeing outcomes – Physical aggression was measured using six items from the Child Behavior Checklist aggression sub scale (Achenbach 1991; foster parent completed) and a measure compiled using seven items from the Sutter–Eyberg Student Behavior Inventory—Revised (Eyberg and Pincus, 1991; teacher completed). Self-control was measured using a scale (Wills et al. 2007) completed by both foster parents and teachers (parallel versions). NB Statistical analyses of teacher ratings are not reported as these showed no change. Satisfaction with services – Measured using a questionnaire – statistical analysis not presented.		

13. Lind T, Bernard K, Ross E et al. (2014) Intervention effects on negative affect of CPS-referred children: Results of a randomized clinical trial. Child Abuse and Neglect 38: 1459–67

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim – The study aimed to determine the effectiveness of Attachment and Biobehavioral Catch-up, an intervention designed to enhance children's self-regulatory capabilities, for young children who had been referred to Child Protective Services. The authors hypothesised that children who participated in the Attachment and Biobehavioral Catch-up programme (with their parents) would display lower levels of negative affect whilst participating in a challenging task than those who participated in the control intervention (with their parents). **Methodology** RCT **Country** Not UK	**Participants** Children and young people. – Children who had been referred to Child Protective Services as a result of maltreatment. Eligibility criteria were: age of less than 2 years at the point of referral to Child Protective Services and residence with their biological parent. **Caregivers and families** – Biological parents of children referred to Child Protective Services as a result of maltreatment. **Sample characteristics** Age – NB. Demographic data collected at post-intervention follow-up. Age of child at post-intervention follow-up: Intervention M=26.7 months (SD=3.8), control M=26.2 months (SD=3.0). Age of caregiver at post-intervention follow-up: Intervention M=28.7 years (SD=7.5), control M=27.7 years (SD=8.3). Sex – Sex - Child - Intervention n=31 male (55%), n=25 female (45%). Control n=31 male (51%), n=30 female (49%).	**Effect sizes** **Children and young people's health and wellbeing outcomes** – Anger: Children in the intervention group showed significantly lower levels of anger (small to medium effect size) than those in the comparison group: F (1, 115)=4.69, p<.05, d=0.40. Anger toward parent: Children in the intervention group showed significantly lower levels of anger towards parent (small to medium effect size) than those in the comparison group: F (1, 115)=5.35, p< 05, d=0.43. Global sadness/anger: Children in the intervention group showed significantly lower levels of global sadness/anger (small to medium effect size) than those in the comparison group: F (1, 115)=5.66, p< 05, d=0.44. Composite negative affect score: Children in the intervention group showed significantly lower levels of affect expression (small to medium effect size) than those in the comparison group: F (1, 115)=5.04, p< 05, d=0.42. **Narrative findings - effectiveness** – Children in the intervention group showed significantly lower scores on the anger, anger toward parent, and the global sadness/anger scales of the Revised Manual for Scoring Mother Variables in the Tool-Use Task as well as a composite measure of negative affect devised from these three scales. All differences showed a small to medium effect size.	**Overall assessment of internal validity** - **Overall assessment of external validity** ++ **Overall validity score** - The decision to use the Tool Task to observe the child's emotional expression and the use of an unpublished scale to score these (particularly without explanation) and the failure to provide detail on methodological issues such as exposure and contamination mean that it is difficult to be confident in the authors findings.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
– USA ('... a large, mid-Atlantic city ...' (p1461). **Source of funding** Government. – National Institute of Mental Health.	Caregiver - Intervention n=2 male (4%), n=52 female (96%). Control n=1 male (2%), n=57 female (98%). Ethnicity – Child - Intervention n=5 White (9%), n=35 African American (62%), n= 14 Biracial (25%), n=2 Hispanic (4%). Control n=5 White (8%), n=37 African American (61%), n= 5 Biracial (8%), n=14 Hispanic (23%). Caregiver - Intervention n=10 White (19%), n=35 African American (65%), n=4 Biracial (7%), n=5 Hispanic (9%). Control n=6 White (10%), n=36 African American (62%), n=1 Biracial (2%), n=15 Hispanic (26%). Religion/belief – Not reported. Disability – Not reported. Long term health condition – Not reported. Sexual orientation – Not reported. Socioeconomic position – Socioeconomic position - Parent level of education (from n=103 parents who provided information): Failed to complete high school 68%, high school diploma 29%, '… completed some college …' (p1461) 3%. Type of abuse – Families had been referred to Child Protective Services for allegations of		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	maltreatment, however the authors did not have access to case records and were unable to determine specific reasons for referral. They also note that substantiation status was not a criterion for eligibility. Looked after or adopted status – Eligibility criteria for the study included residence with biological parents. After enrolment a number of children were removed from the care of their families (n=12). Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size** Comparison numbers – n=131 initially randomised to comparison condition, follow-up data available for n=61 children. Intervention number – n=129 initially randomised to experimental condition, follow-up data available for n=56 children. Sample size – n=260 initially randomised (control n=131, intervention n=129), final sample for whom data were available n=117. **Intervention category** - Attachment and Biobehavioral Catch-up.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Intervention **Describe intervention** – Attachment and Biobehavioural Catch-up (manualised) - An intervention designed to help parents to behave in ways which support their child's self-regulation skills in relation to affect, behaviour and physiology. **Delivered by** – Parent coaches (with some supervision) who had '... strong interpersonal skills and past experience working with children ...' (p1462) and a mixture of bachelor and masters level education. **Delivered to** – Biological parent/child dyads referred to Child Protective Services for allegations of maltreatment. (The majority of participating parents were female). Eligibility criteria were: child must reside with their biological parents and be under the age of two at referral to Child Protective Services. **Duration, frequency, intensity** – 10 sessions (no further details are provided). **Key components and objectives of intervention**		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	– The key objectives of the intervention were to '…enhance children's ability to develop secure and organised attachments, to develop normal cortisol production, and to develop the ability to regulate emotions effectively …' (p1462). To achieve this the intervention is intended to change parenting behaviours to ensure that responses to the child are: 1. Synchronous (i.e. following the child's lead and giving them control, and responding quickly and with sensitivity. It is suggested that this improves regulation of affect). 2. Nurturing (i.e. responding sensitively to distress which is hypothesised to enable children to manage negative affect and to develop secure and organised attachments). 3. Non-frightening (The authors give examples of frightening behaviour such as rough play and a failure to respond to cues from the child, physically intrusive behaviour, shouting or hitting, etc. They note that frightening behaviour is associated with disorganised attachment and difficulties in the control of emotion expression.) The authors note that maltreatment in children often leads to difficulties in the regulation of emotion (particularly during challenging situations) which can be exacerbated by the failure of maltreating parents to interact with their child in ways which		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	help them to develop self-regulation skills. Although the intervention is based on content outlined in a manual the authors note that the '… parent coach's primary role was to provide 'in the moment' feedback about the parent's interaction with his or her child.' (p1462) This feedback focused on the three concepts of synchrony, nurturance and non-frightening behaviour and was intended to help parents to understand and practice these target behaviours and to gain an awareness of how their behaviour can impact upon the development of their child. **Content/session titles** – Sessions 1 and 2 focused on nurturing responses to child distress. Sessions 3 and 4 focused on synchronous responses. Sessions 5 and 6 focus on avoiding the use of frightening and intrusive behaviours. Sessions 7–10 were tailored to the needs of each parent. (Sessions 7 and 8 included discussion of how the parents own attachment histories might impact upon the way in which they interacted with their child. Although each session had a specific topic the target behaviours were emphasised throughout all sessions through the coaches use of 'in the moment' feedback.)		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Location/place of delivery – Family home. **Describe comparison intervention** – Developmental Education for Families (manualised) - Adapted from a home-visiting programme, previously shown to be effective in improving intellectual functioning. The authors note that the components of this intervention which addressed parental sensitivity were removed in order to ensure the two conditions remained distinct. **Delivered by** Parent coaches (with some supervision) who had '... strong interpersonal skills and past experience working with children ...' (p1462) and a mixture of bachelor and master's level education. **Delivered to** Biological parent/child dyads (The majority of parents were female). **Duration, frequency, intensity** Ten sessions (frequency and length not reported although authors note the two conditions were comparable in relation to these aspects). **Key components**		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	The programme is designed to strengthen children's cognitive, motor and language skills. **Content/session titles** - N/A. Delivered in the family home. **Outcomes measured** Children and young people's health and wellbeing outcomes – Children's expression of emotion was assessed during a parent-child problem-solving activity known as the Tool Task (Matas et al. 1978). This consists of 3 increasingly difficult tasks and the final 2 are impossible for the child to complete without help. At the beginning of the session, parents were told that the tasks were too difficult for the majority of young children to complete themselves. They were then instructed to let their child attempt to solve the problem on their own for a few minutes and then offer whatever help they thought necessary. A composite score of negative affect was devised using three scales outlined in the Revised Manual for Scoring Mother Variables in the Tool-Use Task (Sroufe et al., 1980 Revised manual for scoring mother variables in the tool-use task. University of Minnesota: Unpublished). These scales were: anger,		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	anger toward parent and global sadness/anger. Scores on these scales are also reported separately.		

14. Mast JE, Antonini TN, Raj SP et al. (2014) Web-based parenting skills to reduce behavior problems following abusive head trauma: A pilot study. Child Abuse and Neglect 38: 1487–95

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim – The study aimed to examine the '... efficacy of a web-based intervention with live coaching designed to improve parenting skills and everyday child functioning ...' (p1488) for children who had experienced abusive head trauma. **Methodology** RCT. **Country** Not UK. – Study conducted in the USA. **Source of funding** Government.	**Participants** **Children and young people** – Children between the ages of 3 and 8 who had suffered abusive head trauma. Inclusion criteria for the larger trial from which the sample were drawn were – Score of 12 or less on the Glasgow Coma Scale or evidence of brain injury using computerised tomography or magnetic resonance imaging; head trauma had required hospitalisation; trauma had occurred after birth; child between the ages of three and eight at enrolment; inpatient rehabilitation (if needed) had ceased; child remained with legal guardian for duration of study; families should speak English as their primary language. Exclusion criteria were – Parent or child with a history of hospitalisation as a result of psychiatric diagnosis; or the child had a developmental disability or	**Effect sizes** **Quality of parenting and parent-child relationships** – Parent-child interactions (parenting skills and child compliance) were measured using the Dyadic Parent-Child Interaction Coding System. Videotaped play sessions (child and parent led) were coded. Coding focused on positive parenting behaviours such as reflective statements and behavioural descriptions) and undesirable parenting behaviours such as (questions, criticisms, or commands). NB Parental descriptions of child behaviour were removed from analyses as a result of low reliability. Child compliance and 'parent follow-through' were also coded using the following variables: 'Percentage of times that child complied with parent direct commands' (p1491). 'Percentage of times that parent gave child labelled praise after child followed direct command' (p1491). Percentage of times parent properly used discipline techniques when child did not comply with direct command' (p1491). Child directed interactions at follow-up (Primary and secondary caregiver data was averaged) – Parents' use of labelled praise: Parents in the intervention group were significantly more likely to use labelled	**Overall assessment of internal validity** - **Overall assessment of external validity** ++ **Overall validity score** - Key study limitations: Very small sample size (n=9), and resulting very low level of statistical power (12 to 22% at 0.05 criterion - usual standard would be 80%).

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
– National Institute on Disability and Rehabilitation Research.	significant intellectual difficulties which were not a result of the brain injury. **Caregivers and families** – Families of children who had suffered abusive head trauma. Assessments appear to have been conducted with both primary and secondary caregivers in some instances but it is not specified whether the intervention was delivered to both caregivers. **Sample characteristics** Age – No details provided, only that children were between the ages of 3 and 9 at enrolment. Sex – Child - not reported. Primary caregiver - Intervention n=3 female, n=1 male. Control n=2 female, n=1 male. (NB not reported for families who dropped out or were lost to follow up.) Ethnicity – Intervention - Caucasian n=4, African American/multiracial n=0. Control - Caucasian n=2, African American/multiracial n=1. (NB not reported for families who dropped out or were lost to follow up and not specified whether this corresponds to child or primary caregiver.) Religion/belief – Not reported.	praise than those in the comparison group, (reported as both p=.027 and p=.03). No relative risk values given. Parents' use of reflective statements: Parents in the intervention group were significantly more likely to use reflective statements than those in the comparison group, relative risk=9.35, lower relative risk=3.91, upper relative risk=22.39, chi-square = 25.18, p<.001. Parents' use of questions: Parents in the intervention group were significantly less likely to ask their child questions than those in the comparison group, relative risk=0.31, lower relative risk=0.20, upper relative risk=0.50, chi-square=2.052, p<.001. Parents' use of commands: No significant difference between groups, relative risk=0.66, lower relative risk=0.37, upper relative risk=1.17, chi-square=2.041, p=.153. Parent directed interactions at follow-up - Primary and secondary caregiver data was averaged) – Parents' use of labelled praise: Parents in the intervention group were significantly more likely to use labelled praise than those in the comparison group, relative risk=16.9, lower relative risk=2.25, upper relative risk=127.7, chi-square=7.58, p=.006. Parents' use of reflective statements: Parents in the intervention group were significantly more likely to use reflective statements than those in the comparison group, relative risk=13.9, lower relative risk=4.21, upper relative risk=46.19, chi-square=18.60, p<.001. Child compliance following direct commands (%): Significant difference, in favour of the intervention group, estimate=0.39, standard error=0.11, t=3.60, p=.023 (also reported as p=.02).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Disability – Not reported. Long term health condition – All children had a lowest recorded score of 12 or less on the Glasgow Coma Scale or showed evidence of brain injury using computer tomography or magnetic resonance imaging (eligibility criteria for larger trial from which sample were identified). Sexual orientation – Not reported. Socioeconomic position – Marital status of primary caregiver - Intervention n=3 married or living with someone, n=1 not married/divorced. Control n=2 married or living with someone, n=1 not married/divorced. (NB not reported for families who dropped out or were lost to follow up.) Education level of primary caregiver - Intervention n=2 high school diploma or less, n=2 2 or more years of college. Control n=1 high school diploma or less, n=2 2 or more years of college. (NB not reported for families who dropped out or were lost to follow up.) Employment status of primary caregiver - Intervention n=3 working full or part time, n=1 not working. Control n=1 working full or part time, n=2 not working. (NB not reported for families who dropped out or were lost to follow	Parental labelled praise following child compliance (%): No significant difference between groups, estimate = 0.17, standard error=0.08, t=2.03, p=.099. Parental use of clear consequences following child noncompliance for a direct command (%): The authors report that no parents in either group provided clear consequences for noncompliance in response to a direct command. **Children and young people's health and wellbeing outcomes** – Parental ratings of child behaviour - Eyberg Child Behavior Inventory - Total Intensity at follow-up (controlling for baseline scores): Children in the intervention group scored significantly lower than those in the comparison group, with a very large effect size - F(1,4) =13.07, df=2, p=.02, partial eta-squared=.77. Child behaviour problems - Eyberg Child Behavior Inventory - Total problems at follow-up (controlling for baseline scores): No significant difference between groups, F=0.01, df=2, p=.91, partial eta-squared=0. Child behaviour - Child Behavior Checklist – Internalising problems after controlling for baseline scores): No significant difference between groups, F=0.26, df=2, p=.64, partial eta-squared=.65. Child behaviour - Child Behavior Checklist - Externalizing problems after controlling for baseline scores): No significant difference between groups, F=0.31, df=2, p = .61, partial eta-squared=0.7. Child behaviour - Child Behavior Checklist - Total problems after controlling for baseline scores): No significant difference between groups, F=0.10, df=2, p=.76, partial eta-squared = .03.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	up.) Home computer at baseline - Intervention n=1 had a home computer, n=3 did not have a home computer. Control n=1 had a home computer, n=2 did not have a home computer. (NB not reported for families who dropped out or were lost to follow up.) Type of abuse – Physical. All children had an acquired brain injury which determined to be the result of abuse by a hospital multidisciplinary child abuse team. Looked after or adopted status – Two children were living with adoptive parents. The authors also note that five children were living with parents NOT suspected of being the abuser, whilst two children lived with a caregiver who WAS suspected of causing the injury. Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size** Comparison numbers – n=4 Intervention number – n=5 **Sample size** – n=9 **Intervention category** – I-InTERACT (Internet-based Interacting Together Everyday: Recovery	**Narrative findings - effectiveness** **– Quality of parenting and parent-child relationships** – Parents in the intervention group were significantly more likely to use labelled praise and reflective statements; and significantly less likely to use questions than parents in the comparison group during child directed interactions. No significant differences between groups were found in parental use of commands during child-directed play. No parents in either group provided clear consequences for noncompliance in response to a direct command. Parents in the intervention group were also significantly more likely to use labelled praise and reflective statements in parent directed interactions than those in the comparison group. Children in the intervention group were significantly more likely to comply with their parents' commands during parent directed interactions than children in the comparison group. There were no significant differences between groups in parental use of labelled praise following child compliance during parent directed interactions. **Children and young people's health and wellbeing** **-** Parental ratings of child behaviour: Children in the intervention group had significantly lower scores on the total intensity scale of the Eyberg Child Behavior Inventory, with a very large effect size; however there were no significant differences between groups in scores on the total problems scale of this measure. There were also no significant differences between groups in scores on the internalizing, externalising, or total problems scale of the Child Behavior Checklist.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	After Childhood Traumatic Brain Injury). Manualised. **Intervention** **Describe intervention** – I-InTERACT (Internet-based Interacting Together Everyday: Recovery After Childhood Traumatic Brain Injury). **Delivered by** – 'Three Master's level research personnel (a research coordinator and two advanced clinical psychology doctoral students)' (p1490). The three 'therapists' received training in the sequelae of traumatic brain injury and were instructed on how to deliver the intervention. They also had weekly supervision meetings with a licensed clinical psychologist. **Delivered to** – Parents of children with abusive head trauma. Some of whom may have been suspected perpetrator. **Duration, frequency, intensity** – Ten core sessions (with the option for supplementary sessions delivered between core sessions 9 and 10. Families could schedule four supplementary sessions from a range of five topics). Sessions 1–9 were intended to be	**Narrative findings - qualitative and views and evidence** The authors report that families valued the intervention and the behaviour management techniques which they had been able to develop as a result, as well as the opportunity to get direct feedback on their parenting. NB The qualitative findings are quite brief and it is not clear which Participants provided this feedback (e.g. whether feedback was provided from Participants who did not complete the intervention).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	completed within 3 months, with months four and five used to deliver supplementary sessions where families requested these. The tenth core session was completed in month 6. Families who did not schedule supplementary sessions were contacted on a bi-weekly basis between the core sessions 9 and 10 via phone or email (four contacts in total) to discuss their progress. **Key components and objectives of intervention** – The I-InTERACT programme is a parenting skills programme which focuses on positive parenting skills and consistent use of discipline. It incorporates content from a number of parenting programmes including Parent-Child Interaction Therapy but also helps parents to develop behaviour management techniques to address the difficulties that children who have experienced head injuries may have in learning from consequences. The programme is based on therapy protocol outlined in a manual. The programme also includes content on the behavioural sequelae of head injury as well as communication issues and management of stress. Optional sessions focus on specific ongoing problems		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	which the families may be experiencing such as pain management or guilt. **Content/session titles** – Session 1 'Introduction to I-InTER-ACT Program'; Session 2 'Introduction to Positive Parenting Skills and Special Play Time'; Session 3 'Staying Positive and Coping with Stress'; Session 4 'Behavior Management'; Session 5 'Introduction to 'Lead Your Child' (Parent Directed Interaction)'; Session 6 'Dealing with Anger'; Session 7 'Introduction to Consequences for Not Following Directions'; Session 8 'Cognitive Problems'; Session 9 'House Rules and Using Positive Parenting Skills in Real Life'; Session 10 'Closing Thoughts' (p1490). Titles of optional sessions are as follows: 'Marital Communication'; 'Parents and Siblings'; 'Pain Management'; 'Guilt and Grief'; 'Working with the School and Transition Issues' (p1490). **Location/place of delivery** – Sessions are conducted in the family home. The first session is delivered during a home visit but the remainder are delivered online and take the form of a web module (including reading about specific skills, watching videos of parents demonstrating these skills, and completing exercises on these		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	skills) and a videoconference session via Skype or Movi Client during which the parents and therapists review the web module, role-play the skills learnt during the web module. The parent then plays with their child and practices these skills whilst receiving 'bug-in-the-ear' feedback from the therapist. Supplementary sessions included both a web module and videoconference session. Families also appear to have been given access to a website providing information on abusive head trauma. **Describe comparison intervention** – Internet Resource Comparison Group. Families in this group were given access to a study website that provided links to a range of internet resources which included information on traumatic brain injury, support groups and associations, as well as resources focusing on recovery, coping and parenting skills. Families were asked to use these as often as they liked and to note which they found the most useful. **Outcomes measured** Quality of parenting and parent-child relationships – Parent-child interactions (changes in parenting skills and child compliance from pre to post-test) were assessed		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	using the Dyadic Parent-Child Interaction Coding System (Eyberg, Nelson, Duke et al. 2005). Videotaped play sessions (child and parent led) were coded. Coding focused on positive parenting behaviours such as reflective statements and behavioural descriptions) and undesirable parenting behaviours such as (questions, criticisms, or commands). NB Parental descriptions of child behaviour were removed from analyses as a result of low reliability. Child compliance and 'parent follow-through' were also coded using the following variables: 'Percentage of times that child complied with parent direct commands' (p1491). 'Percentage of times that parent gave child labelled praise after child followed direct command' (p1491). 'Percentage of times parent properly used discipline techniques when child did not comply with direct command' (p1491). Children and young people's health and wellbeing outcomes – Child behaviour measured using Eyberg Child Behavior Inventory (Eyberg and Pincus 1999; parental report - measures frequency and intensity of a child's oppositional behaviour and conduct problems - provides total problems and total intensity scores. Scores over 65 are considered to be clinically		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	elevated.) Child behaviour was also measured using the Child Behavior Checklist (Achenbach and Rescorla 2000, 2001; parental report - 1 of 2 versions used depending on age of child - measures behavioural and emotional functioning over last 6 months - authors used composite scales measuring internalizing behaviours, externalizing behaviours and total problems). Scores of 65 or more are considered to be clinically elevated.		

15. Oxford ML, Fleming CB, Nelson EM et al. (2013) Randomized trial of Promoting First Relationships: Effects on maltreated toddlers' separation distress and sleep regulation after reunification. Children and Youth Services Review 35: 1988–92

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim – The study aimed to determine the impact of Promoting First Relationships, an attachment focused intervention, for toddlers recently reunified with their biological parent after being placed in foster care. The study specifically evaluated the impact of the inter-	**Participants** Children and young people – Toddlers with a change in primary caregiver in the previous 7 weeks as a result of a court-order, recently reunified with their biological parent. (The age of the participating children is variously reported as ten to 24 months and 11 to 36 months). **Caregivers and families** – Biological parents of toddlers who had experienced a court-ordered	***Effect sizes*** Children and young people's health and wellbeing outcomes – NB Dyads were excluded from analysis if the child was removed from the parents care during the course of the study. Data therefore based on 43 dyads (control – n=25, intervention – n=18) Sleep problems at six month follow-up: Comparison of means showed no significant difference between groups (t=1.54, p=.132). However, regression analysis adjusting for covariates including sleep problems at baseline showed that being in PFR was a significant predictor of reduced sleep problems (b=-2.116,	**Overall assessment of internal validity** - **Overall assessment of external validity** ++ **Overall validity score** - Key study limitations: High attrition rate, particularly in intervention group. The exclusion from the study of dyads that were no longer in

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
vention on sleep problems. The authors decided to analyse the subsample of birth parents enrolled in the original study due to their increased risk and significant demographic differences when compared to foster carers or kinship carers participating in the parent study. The authors also note that it is more likely that this group will experience feelings of greater anxiety and inadequacy in the parenting role as a result of the removal of their child from their care and would therefore derive greater benefit from the intervention. The authors hypothesise that the intervention helps to develop parent's ability to respond sensitively to the behavioural cues of their child and support their ability to self-regulate, thereby leading to reductions in separation	change in primary caregiver in the previous seven weeks and had recently been reunified with their child. **Sample characteristics** Age – Age - Infant mean age in months (standard deviation) - Control 18.15 (4.79), intervention 18.29 (5.32). Age of caregivers not reported. Sex – Male target child – Control n=11 (44%), intervention n=9 (50%). Male caregiver – Control n=4 (16%), intervention n=1 (6%). Ethnicity – Infant - Control - Hispanic n=3 (12%), Native American/Alaskan native n=1 (4%), Black n=1 (4%), mixed race n=2 (8%), 'unable to determine' n=2 (4%), White n=19 (76%). Intervention - Hispanic n=0 (0%), Native American/Alaskan native n=0 (0%), Black n=3 (17%), mixed race n=5 (28%), 'unable to determine' n=0 (0%), White n=10 (56%). The study required eligible caregivers to speak English. Religion/belief – Not reported. Disability – Not reported. Long term health condition – Not reported. Sexual orientation	p<0.05). (NB p value is denoted with *, but this is missing from the table key. However, tables elsewhere in the paper use * to denote a value of <0.05). The effect size, in terms of standard deviation unit difference, was d=0.67. Separation distress at 6 month follow-up: Comparison of means showed significant difference between groups in favour of the intervention - (t=3.05, p=.004). Study had predicted that sleep problems would be mediated by separation distress. This was tested using a path model. This showed a significant association between the intervention and reduced separation distress (path coefficient = -0.45, p<0.05), and a significant association between reduced separation distress and reduced sleep problems (path coefficient = -0.34, p<0.05). The unstandardised estimate for the indirect effect of PFR on sleep problems (mediated by separation distress) was -0.96 (95% CI =2.22 to -0.07). **Narrative findings - effectiveness** – The study found that being in the intervention group (compared to control group) significantly predicted reduction in sleep problems, with medium to large effect size (d=0.67). Path analysis showed that the relationship between being in the intervention group and reduced sleep problems, was mediated by impact on separation distress.	the same household (presumably those in which the child had been removed back in to care) is a possible source of bias, as these are likely to be families with the highest level of need, for whom the intervention may have been less likely to be effective. Exclusion of these families may therefore have inflated estimates of the effectiveness of the intervention. The sample size for the study is also relatively small (n=43), and there is no consideration in the paper of study power.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
distress and sleep problems. **Methodology** – Secondary analysis of data from an RCT focusing on a subsample (biological parent-child dyads) of the overall population (included kinship carers and foster carers) originally randomised (reported in Spieker et al. 2012). **Country** Not UK. – USA (single county). **Source of funding** Government. – National Institute of Mental Health.	– Not reported. Socioeconomic position – Household income of less than $20,000 per year – control n=14 (58%), intervention n=12 (71%). Type of abuse – Not reported. Looked after or adopted status – Removed from birth parents home on more than one occasion – control n=1 (4%), intervention n=5 (28%). Mean age of child in months at first removal (standard deviation) – control 8.59 (6.78), intervention 7.23 (6.86). Mean number of changes in caregivers prior to enrolment – Control 3.04 (1. 14), intervention 2.94 (1.16). Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size** Comparison numbers – n=29 randomised. NB Only data from dyads which were remained intact were used in analyses. At the six month assessment only 25 remained intact. Intervention number – n=27 randomised. NB Only data from dyads which were remained intact were used in analyses. At the six month assessment only 18 remained intact.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Sample size – n=56 randomised. NB Only data from dyads which remained intact were used in analyses. At the six month assessment only 43 remained intact. **Intervention category** Parenting programmes. **Intervention** **Describe intervention** – Promoting First Relationships. **Delivered by** – Providers from community mental health services. **Delivered to** – Biological parents of children between the ages of 10 and 24 months whose child had experienced a court-ordered change in primary caregiver in the last seven weeks but had recently been reunified with their parent. **Duration, frequency, intensity** – Ten weekly sessions around 60–75 minutes in length. The authors report that 67% of parents in the intervention group took part in all ten sessions. **Key components and objectives of intervention**		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	– The authors report that Promoting First Relationships is designed to '…improve sensitive, responsive, and predictable care by caregivers (foster, kin, and birth caregivers) of toddlers with a recent child welfare mandated placement change' (p1989). One session focused on caregiver responses to separation distress, and responses to distress more generally are reported as a wider theme of the intervention. Work focusing specifically on separation distress involved the use of a videotaped separation and reunion between the caregiver and child which the provider used to help the caregiver to understand the emotional cues the child showed (e.g. indirect displays of distress, or inconsolable distress which the caregiver might interpret as manipulative, etc.) and recognise the need for a predictable relationship with their child. This work also involved encouraging the parent to reflect on their recent separation and reunification with their child and addressed their feelings of anger, anxiety and guilt regarding these events. **Content/session titles** – N/A. **Location/place of delivery** – Caregiver's home.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Describe comparison intervention – Early Educational Services. **Delivered by** Specialists in early education. **Delivered to** Biological parents of children between the ages of ten and 24 months whose child had experienced a court-ordered change in primary caregiver in the last seven weeks but had recently been re-unified with their parent. **Duration, frequency, intensity** Three monthly sessions of 90 minutes; 96% of parents in the comparison group took part in all three sessions. **Key components and objectives of intervention** – Education in early development of children and referral to other services when necessary. **Content/session titles** N/A. **Location/place of delivery** Caregiver's home. **Outcomes measured** Children and young people's health and wellbeing outcomes		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	– Parental reports of sleep problems were measured using 4 items from the Child Behavior Checklist (Achenbach and Rescorla, 2000) and 2 items from the Brief Infant Toddler Social and Emotional Assessment (Briggs-Gowan & Carer 2002). The score on this scale was the average of the 6 items. NB Only measured at the 6 months post-intervention assessment due to the age of participating children at base-line and immediate post-intervention assessment. Separation distress was measured using the Separation Dis-tress meaning cluster of the Toddler Attachment Sort-45 (Kirkland et al. 2004), a modified version of the At-tachment Q-Sort (Waters 1987). The authors report that this cluster '… weights most heavily on the following items: 'When mom talks with others, child wants attention', 'Child is very clingy, stays close to mom', 'Child cries when mom leaves or moves to another place, and Child gets upset if mom leaves or shifts place' (p1991). NB This was measured during a home visit in which researchers observed the child's distress when their parent ap-peared to leave the home.		

16. Purvis KB, Razuri EB, Howard ARH et al. (2015) Decrease in Behavioral Problems and Trauma Symptoms Among At-Risk Adopted Children Following Trauma-Informed Parent Training Intervention. Journal of Child and Adolescent Trauma 8: 201–210

Research aims.	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating.
Study aim Study aim: 'Using a two-group, pre-post intervention design, the current study evaluated the effectiveness of a parent training utilizing Trust-Based Relational Intervention, a trauma-informed, attachment-based intervention, in reducing behavioral problems and trauma symptoms in at-risk adopted children' (p201). **Methodology** RCT including cluster. **Country** Not UK. USA.	**Participants** • Children and young people. • Caregivers and families. **Sample characteristics** • Age - Age in years: Control group (7.88) Intervention group (7.88) • Sex - Intervention: 30 male 18 female Control: 30 male 18 female • Ethnicity - Intervention: Asian 9 Black/African American 13 Hispanic/Latino 1 White/Caucasian 20 Native American 1 Other 0 Control: Asian 12 Black/African American 16 Hispanic/Latino 4 White/Caucasian 16 Native American 0 Other 0 • Type of abuse – Intervention: Neglect Yes 35 No 13; Physical abuse Yes 13 No 35; Sexual abuse Yes 7 No 41. Control: Neglect Yes 37 No 11; Physical abuse Yes 19 No 29; Sexual abuse Yes 7 No 41. • Looked after or adopted status - All participants were adopted. **Sample size** • Comparison numbers - 48 children • Intervention numbers - 48 children **Describe intervention** 'TBRI is a trauma-informed intervention grounded in attachment theory	**Effect sizes** Children and young people's health and wellbeing outcomes. The study found significant effects in relation to four of the five SDQ subscales (as reported by care-givers). Care-givers reported that emotional problems, conduct problems, hyperactivity/inattention and total difficulties were 'significantly lower' following the intervention. The results in the post-test found an improvement in children who had had the intervention. There was no significant improvement over time for children in the control group. Significant interaction between time and age (current) Emotional Problems: $p<.05$ $\eta p 2=.05$), Peer Problems: $p<.01$, $\eta p2=.08$), Total Difficulties: ($p<.05$, $\eta p2=.06$).	**Overall assessment of internal validity** + **Overall assessment of external validity** ++ **Overall validity score** +

Research aims.	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating.
	that seeks to improve outcomes for vulnerable children by (1) helping caregivers see the needs of children who have experienced relational trauma and (2) helping caregivers do what is necessary to meet those needs ... TBRI consists of three sets of principles that facilitate felt-safety, self-regulation, and connection: Empowering Principles, Connecting Principles, and Correcting Principles. Each set of principles has two associated sets of strategies' (p203). **Delivered by** Trainers who had a least two years experience of TBRI parent training. **Delivered to** Adoptive parents of children who had a history of maltreatment. **Duration, frequency, intensity, etc.** The treatment group attended a four day TBRI training session designed to teach strategies and skills to improve behaviour. **Key components and objectives of intervention** The TBRI intervention is based on three principles that seek to improve outcomes for vulnerable children. The principle are: Empowering principles -		

Research aims.	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating.
	Aiming to help care givers develop the child's capacity for self-regulation and decrease negative and disruptive behaviour. Connecting principles - Aiming to help build trusting relationships and connect the other two principles Correcting principles - To shape behaviour and responses. Each of these principles is linked to strategies. Empowering principles: Ecological strategies and psychological strategies. Connecting principles: Mindful awareness and engagement strategies. Correcting Principles: Proactive strategies and responsive strategies. **Content/session titles** On site group: Four-day TBRI parent training. One day on each principle and an overview day. The days consisted of lectures, standardised presentations, group discussions, therapy, group activities and videos. **Location/place of delivery** Four training sessions at the university. **Describe comparison intervention** The control group were offered on-line training. **Outcomes measured** Children and young people's health		

Research aims.	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating.
	and wellbeing outcomes The assessments: Strengths and difficulties questionnaire and the trauma symptom checklist. Strengths and difficulties measures behavioural problems Trauma Symptom checklist measures posttraumatic symptoms The study reports on short-term improvements in behaviour and trauma		

17. Reddy SD, Negi LT, Dodson-Lavelle B et al. (2013) Cognitive-Based Compassion Training: A promising prevention strategy for at-risk adolescents. Journal of Child and Family Studies 22: 219–30

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim – Aim of study to examine whether a 6-week Cognitive-Based Compassion Training (CBCT) intervention would improve psychosocial functioning among adolescents in foster care. **Methodology** RCT. **Country** Not UK. – US study.	**Participants** **Children and young people** – A sample of children aged 13–17 living in the foster care system. **Sample characteristics** Age – Mean age 14.7 (sd=1.14) Sex – Female 56%, Male 44% Ethnicity – 78.8% African American Religion/belief – Not reported Disability – Not reported Long term health condition	**Effect sizes** **Children and young people's health and wellbeing outcomes** – Depressive symptoms measured using the Quick Inventory of Depressive Symptomatology-Self report (QIDS-SR; Rush et al. 2003): No significant differences between groups post-treatment (no data provided). Anxiety measured using the State-Trait Anxiety Inventory (Spielberger et al. 1983). No significant differences between groups post-treatment (no data provided). Self-harm measured using functional assessment of self-mutilation (FASM, Lloyd et al. 1997): No outcome data reported (no data provided).	**Overall assessment of internal validity** - **Overall assessment of external validity** + Unable to award ++ due to lack of specifity regarding maltreatment history of Participants. **Overall validity score** - Poor reporting of sample size and attrition rates. Unclear whether assessors were blinded

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	– Psychiatric disorder: One axis I disorder 37%; More than 1 axis 1 disorder 51% (anxiety disorders 6%, depression/dysthymia 36%, attention-deficit hyperactivity disorder 40%, oppositional defiant/conduct disorder 43%, adjustment disorder 10%, bipolar disorder 7%, post-traumatic stress disorder 10%). One psychiatric medication 13%, two or more psychiatric medications 29%. BMI: 20% overweight (BMI 25-29.9) 11.4% obese (BMI of 30 or above). Sexual orientation – Not reported. Socioeconomic position – Not reported. Type of abuse – Experiences of abuse measured using the Child Trauma Questionnaire (Fink et al. 1995). Scale measuring abuse ranging from 0 to 20 (0 to 12 for minimisation/denial scale), with higher scores indicating more abuse. Scores (standard deviations) for CBCT group; waitlist Emotional abuse: 5.92 (6.0); 5.56 (5.12) Physical abuse: 6.83 (6.81); 4.29 (5.23) Sexual abuse: 2.72 (5.20); 3.26 (5.91) Emotional neglect: 6.69 (5.11); 7.50 (5.80) Physical neglect: 4.03 (3.97); 3.74 (3.25) Minimisation/denial: 8.42 (3.04); 8.59 (3.28) Looked after or adopted status – All Participants in foster care	Respondent experiences of positive and negative emotions assessed using Self-other Four Immeasurables Scale, SOFI, Kraus and Sears 2009): No between-groups comparisons were conducted. Agency (beliefs about initiating and moving towards goals) measured using Children's Hope Scale (CHS; Snyder et al. 1991): No significant differences between groups post-treatment (no data provided). Difficulties with Emotion Regulation Scale (DERS; Gratz and Roemer 2004): No significant differences between groups post-treatment (no data provided). Callous and unemotional traits measured using Inventory of Callous and Unemotional traits self report version and parent version (Essau et al. 2006): No between-groups comparisons were conducted. **Narrative findings - effectiveness** – There was no significant improvement on self-report measures of psychosocial outcomes for the CBCT group compared to a wait-list control. However, the authors note that 'CBCT was positively evaluated by the majority of this population of at risk adolescent, and most reported they would recommend it to a friend' (p225).	to treatment condition. No consideration of statistical power. Short follow-up time.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size** Comparison numbers – Not reported. Intervention number – Not reported. Sample size – Unclear if sample is 70 (in abstract) or 71 (see Table 1). **Intervention category** Cognitive-based compassion training. **Intervention** **Describe intervention** – Cognitively-Based Compassion Training is described as 'a type of contemplative practice that teaches active contemplation of loving-kindness, empathy and compassion towards loved ones, strangers and enemies (Ozawa-de Silva and Dodson-Lavelle 2011; Ozawa-de Silva et al. in press)' (p220). **Delivered by** – Not reported. **Delivered to** – Children aged 13–17 in foster care. **Duration, frequency, intensity** – Twice-weekly classes for 6 weeks.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Key components and objectives of intervention – The study states that 'Building on basic mindfulness practice, Cognitively-Based Compassion Training (CBCT) employs a variety of cognitive restructuring and affect generating practices with the long-term goal of developing an equanimity of mind that fosters acceptance and understanding of others (Salzberg 2002)' (p220). **Content/session titles** – Not reported. **Location/place of delivery** – Not reported. **Describe comparison intervention** – Waitlist. **Outcomes measured** Children and young people's health and wellbeing outcomes – Depressive symptoms measured using the Quick Inventory of Depressive Symptomatology-Self report (QIDS-SR; Rush et al. 2003). Anxiety measured using the State-Trait Anxiety Inventory (Spielberger et al. 1983). Self-harm measured using functional assessment of self-mutilation (FASM,		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Lloyd et al. 1997) Respondent experiences of positive and negative emotions assessed using Self-other Four Immeasurables Scale, SOFI, Kraus and Sears 2009). Agency (beliefs about imitating and moving towards goals) measured using Children's Hope Scale (CHS; Snyder et al. 1991) Difficulties with Emotion Regulation Scale (DERS; Gratz and Roemer 2004) Callous and unemotional traits measured using Inventory of Callous and Unemotional traits self report version and parent version (Essau et al. 2006). Service outcomes – Post-treatment, Participants completed a 5-item feedback form assessing 1. The helpfulness of the programme, 2. Frequency of thinking about CBCT principles or lessons outside of class, 3. Whether they would recommend the programme to a friend, 4. If they would like to have this programme offered in their schools and 5. How they felt about the length of the programme. The form also included open-ended questions about what they had learned, use of CBCT in daily life and how the programme could be improved.		

18. Rushton A, Monck E, Leese M et al. (2010) Enhancing adoptive parenting: A randomized controlled trial. Clinical Child Psychology and Psychiatry 15: 529–42

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim – The study aimed to evaluate 2 parenting programmes designed to improve adoptive parenting and child behavioural problems. **Methodology** – Two experimental conditions versus care as usual. Due to the small sample size the 2 experimental groups were combined for analysis purposes. **Country** UK – England. **Source of funding** Government – Department of Health and Department for Children, Schools and Families. Voluntary/charity – Nuffield Foundation.	**Participants** **Children and young people** – Adopted children (late placed) between the ages of 3 and 8 assessed as having serious behavioural problems using the Strengths and Difficulties Questionnaire (score must be greater than 13 if this had been completed by the parents or higher than 11 if this had been completed by the child's social worker. **Caregivers and families** – Adoptive parents of children (late placed) between the ages of 3 and 8 assessed as having serious behavioural problems. **Sample characteristics** Age – Childs mean age at placement - Intervention 68 months (SD=19), control 65 months (SD=17). Childs mean age at first admission to care - Intervention 37 months (SD=14), control 27 (SD=7). Sex – Female children - Intervention 53%, control 55%. Ethnicity – Ethnicity (% of white children) - Intervention 84%, control 88%.	**Effect sizes** **Quality of parenting and parent-child relationships** – The authors note that 'due to small sample size the statistical analysis was mainly conducted on the combined interventions versus 'service as usual' cases. At first follow-up assessment (controlled for scores at baseline). Expression of Feelings Questionnaire: Small to medium effect size, non-significant difference between groups - d=0.49, p=0.11. Post Placement Problems: - Very small effect size, non-significant difference between groups - d=0.01, p=0.95. Parenting Sense of Competence Scale - satisfaction with parenting: Small effect size, non-significant difference between groups - d=0.31, p=0.27. Parenting Sense of Competence Scale - parenting efficacy: Small effect size, non-significant difference between groups d=0.20, p=0.46. Daily Hassles (frequency): Small effect size, non-significant difference between groups - d=0.25, p=0.4. Daily Hassles (intensity): Medium effect size, non-significant difference between groups - d=0.53, p=0.09. At second follow-up assessment (controlled for scores at baseline). Expression of Feelings Questionnaire: Small effect size, non-significant difference between groups - d=0.29, p=0.26. Post Placement Problems: Small effect size, non-significant difference between groups - d=0.21, p=0.55. Parenting Sense of Competence Scale - satisfaction with parenting: Medium to large effect size, significant	**Overall assessment of internal validity** - **Overall assessment of external validity** ++ **Overall validity score** - The small sample size and use of a scale with unclear reliability or validity, as well as a lack of detail on key methodological issues such as the use of intent to treat analysis means that it is not possible to award a higher score. The analysis also combines the 2 parenting interventions, making it difficult to draw conclusions about what has led to any improvement in outcomes.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Religion/belief – Not reported. Disability – Not reported. Long term health condition – Not reported. Sexual orientation – Three of the adoptive couples were same-sex. Socioeconomic position – Nine children had been placed with single parents. Severe economic deprivation in birth family - Intervention 52%, control 61%. Type of abuse – Reason for child's first admission to care: Neglect (intervention 89%, control 89%), sexual abuse (intervention 21%, control 22%), physical abuse (intervention 58%, control 44%), emotional abuse (intervention 57%, control 33%), carer's mental illness (intervention 47%, control 39%), carer's addiction (intervention 42%, control 72%), concern about siblings (intervention 56%, control 43%), Schedule 1 offender in household (intervention 16%, control 22%), domestic violence (intervention 63%, control 55%). Looked after or adopted status – Looked after or adopted status - All children had been adopted by non-relative parents. Children placed with existing foster parents were ineligible to	difference between groups in favour of the combined intervention group - d=0.7, p=0.007, 95% CI – 8.4 to -1.4). Parenting Sense of Competence Scale - parenting efficacy: Small effect size, non-significant difference between groups - d=0.34, p=0.21. Daily Hassles (frequency): Very small effect size, non-significant difference between groups - d=0.13, p=0.68. Daily Hassles (intensity): Very small effect size, non-significant difference between groups - d=0.13, p=0.58. The authors also report the analysis of parental management of emotional difficulties and provide statistical analysis of this but this appears to be based on qualitative data with no explanation of how this was collected (e.g. questions asked), as a result this data has not been extracted. **Children and young people's health and wellbeing outcomes** – At first follow-up assessment (controlled for scores at baseline) Total score on the Strengths and Difficulties Questionnaire: Small effect size, non-significant difference between groups - d=0.35, p=0.23. At second follow-up assessment (controlled for scores at baseline). Total score on the Strengths and Difficulties Questionnaire: Very small effect size, non-significant difference between groups - d=0.13, p=0.66. NB. Only data for total scores on this scale are provided which is calculated by adding together scores on the first four subscales which cover emotions, behaviour, restlessness and concentration, and peer relationships. Impact	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	participate. At baseline assessment, the child had been with their adoptive parents for an average of 12 months (range 5 to 18). The mean number of changes in placement prior to this was six for the intervention group (standard deviation 2.9) and 6 for the control group (standard deviation 3.7). Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size** Comparison numbers – n=18 Intervention number – Group 1 (Parent advice type 1 - cognitive behavioural) n=10, group 2 (Parent advice type 2 - educational) n=9. NB Due to the small sample size these two groups were combined for the purposes of analysis. Sample size – n=37. **Intervention category** Parenting programmes – The study evaluated 2 parenting programmes in comparison to care as usual. The first is described as a cognitive behavioural approach and the second is described as an educational approach.	scores and scores on the fifth subscale which measures pro-social behaviour are only reported narratively. Parents perceptions of their child's progress in relation to emotional distress, behaviour and attachment were measured at the final follow-up assessment using a visual analogue scale (mark on a line) however this is only reported in narrative form. Satisfaction with services – Satisfaction with Parenting Advice Questionnaire: No statistical data provided. **Narrative findings - effectiveness** **Quality of parenting and parent-child relationships** At the first follow-up assessment there were no significant differences between groups in scores on the Expression of Feelings Questionnaire; the Post Placement Problems scale; the satisfaction with parenting and the parenting efficacy subscales of the Parenting Sense of Competence Scale; and the frequency and intensity subscales of the Daily Hassles scale. At the second follow-up assessment there was a significant difference in scores on the satisfaction with parenting subscale of the Parenting Sense of Competence Scale which showed a medium to large effect in favour of the intervention group. However, there were no significant differences between groups in scores on the Expression of Feelings Questionnaire; the Post Placement Problems scale; the parenting efficacy subscale of the Parenting Sense of Competence Scale; and the frequency and intensity subscales of the Daily Hassles scale.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Intervention **Describe intervention** – Both the cognitive behavioural and the educational approach were designed to help adoptive parents have greater control of their child's problematic behaviour and to foster a '... consistent, responsive, parenting environment' (p531). **Delivered by** – Both experimental interventions were delivered by children and family social workers who were 'familiar' with adoption (p532). They received training as well as guidance on the use of the manual and could access supervision by the relevant professional who had assisted in the creation of each programme. **Delivered to** – Both experimental interventions were delivered to adoptive parents. In the case of couples, both parents were encouraged to participate. **Duration, frequency, intensity** – Cognitive behavioural approach – 10 sessions (no further information provided). Educational approach – 10 sessions (no further information provided).	**Children and young people's health and wellbeing** – There were no significant differences between groups in total scores, impact scores or scores on any subscales of the Strengths and Difficulties Questionnaire at either the first or second follow-up assessment. The authors also report that on the visual analogue scale of parental judgements of their child's progress (level of emotional distress, misbehaviour, and attachment) there were no significant differences between the groups in their perceptions of the level or direction of change. Satisfaction with services – The authors report that there were no differences between the 2 groups in scores on the Satisfaction with Parental Advice Questionnaire. However, the data for this measure is not reported earlier in the paper. **Narrative findings – qualitative data** – The authors report that adoptive parents who had participated in one of the two parenting programmes '... almost universally responded positively ...' (p538) and note that parents appreciated the ability to work with the same advisor on the specific problems their child had. This was viewed positively in contrast to the generalised advice they had received in the past. Parents who participated in the cognitive behavioural focused group were reported to value the help they received in implementing what they had learnt in a consistent and persistent manner, whilst those in the education focused group appreciated the opportunity to develop their understanding of the situation aided by an experienced practitioner. The advisers involved were reported to value the components of the manuals which focused on play but were less positive about those which focused on aggression. The authors also	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Key components and objectives of intervention – Cognitive behavioural approach (manualised): The cognitive behavioural approach draws on the work of Webster-Stratton and was adapted with the assistance of a clinical psychologist. The programme demonstrates the use of praise as a means of encouraging acceptable behaviour and the use of firm limits, 'logical consequences' and problem-solving as a way of discouraging problematic behaviour. The programme emphasises the importance of daily play sessions and helping parents when their child rejects their praise. The authors report that the intervention includes a cognitive element because '… parenting behaviour is influenced by how adopters construe the child's behaviour and how they come to see themselves in relation to the child' (p531). Educational approach (manualised): The educational approach was designed for the purposes of the study with input from a county adoption adviser. It is intended to enhance the adoptive parents understanding of their child's behaviour and possible triggers of anger or distress. The programme focuses on the adoptive parent's response to	report on parental reports of their handling of emotional difficulties and the way this changed over time. Although, the study includes statistical analysis of this information the NCCSC review team decided not to include this information due to concerns regarding the collection of this data. At the second follow-up assessment parents reported that they managed these types of difficulties differently than they had previously.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	parenting challenges and aims to enable them to anticipate these and consequently be able to manage them better. The advisers who delivered this intervention used local authority adoption records to gain an understanding of each family and the attachment and developmental history of the child. **Content/session titles** – Cognitive behavioural programme - Session 1 – 'Getting to know the parents and introducing the programme', session 2 – 'Using positive attention to change behaviour', session 3 – 'The value of play for establishing positive relationships', session 4 – 'Using verbal praise', session 5 – 'Praise and rewards', session 6 – 'Learning clear commands and boundaries', session 7 – 'Using 'ignoring' to reduce inappropriate behaviour, session 8 – 'Defining for the child the consequences of undesirable behaviour', session 9 – 'Time Out' and problem solving, session 10 – 'Review and ending' (p531). Educational programme - Session 1 – 'Getting to know the parents and introducing the programme', session 2 – 'Understanding insecurity', session 3 – Helping parents understand their own reactions to disturbed children's behaviour, session 4 – 'Understanding how 'bad experiences' affect learning		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	and behaviour', session 5 – 'Understanding how 'bad' and broken relationships affect development', session 6 – 'Children's survival strategies and defensive reactions: the outward show', session 7 – 'The expression and control of feelings', session 8 – 'Understanding how children develop new relationships', session 9 – 'Surviving in the wider world', session 10 – 'Review and ending' (p532). **Location/place of delivery** – Both interventions were delivered in the family home. **Describe comparison intervention** – Services as usual - no further details are provided. **Outcomes measured** Quality of parenting and parent-child relationships – The nature and progress of the parent-child relationship was measured using the Expression of Feelings Questionnaire (Quinton et al. 1998; completed by adopters). This assesses the child's ability to express their feelings and to seek comfort. Higher scores indicate better adjustment. Problematic post-adoption behaviour such as rejection of the new parents was measured using the Post		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Placement Problems questionnaire (completed by adopters). Higher scores indicate more problematic behaviour. Parental satisfaction as well as sense of competence was measured using two scales of the Parenting Sense of Competence Scale (Johnston and Mash 1989; Ohan et al. 2000). Common challenges in the parenting role such as difficulties at mealtimes or arguments between siblings were measured using the frequency and intensity subscales of the Daily Hassles scale (Crnic and Booth 1991). Higher scores indicate more significant problems. The authors also report that they measure parental management of emotional difficulties and provide statistical analysis of this but this appears to be based on qualitative data with no explanation of how this was collected (e.g. questions asked), as a result only the qualitative findings have been extracted from this section. Children and young people's health and wellbeing outcomes – Psychosocial problems were measured using the Strengths and Difficulties Questionnaire (Goodman 2001) which was completed by the adoptive parents. (Only data for total scores on		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	this scale are provided which is calculated by adding together scores on the first four subscales which cover emotions, behaviour, restlessness and concentration, and peer relationships. Impact scores and scores on the fifth subscale which measures pro-social behaviour are only reported narratively.) Parents perceptions of their child's progress in relation to emotional distress, behaviour and attachment were measured at the final follow-up assessment using a visual analogue scale (mark on a line). Satisfaction with services – Parental satisfaction with the 2 experimental interventions were measured using the Satisfaction with Parenting Advice Questionnaire (Davies and Spurr 1998).		

19a. Spieker SJ, Oxford ML, Kelly JF et al. (2014) Promoting First Relationships: Randomized trial of a relationship-based intervention for toddlers in child welfare. Child Maltreatment 17: 271–86

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim – The study aimed to evaluate the impact of the Promoting First Relationships intervention for caregivers of toddlers with a recent placement in foster	**Participants** **Children and young people** – Children between the ages of 10 and 24 months with a change in primary caregiver in the previous seven weeks as a result of a court-order.	**Effect sizes** **Quality of parenting and parent-child relationships** – NB Dyads were excluded from analysis if a change in caregiver occurred during the course of the study. Caregiver outcomes post-intervention (adjusted for baseline score, age of child, multiple removals, caregiver type, and time between baseline and post-intervention assessment). Positive effect sizes represent a	**Overall assessment of internal validity** - **Overall assessment of external validity** ++ **Overall validity score** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
care. The authors hypothesised that the intervention '… would result in improved parenting and child outcomes relative to a comparison condition in which families received home-based services that were not relationship focused' (p273). **Methodology** RCT. **Country** Not UK – USA. **Source of funding** Government – National Institute of Mental Health.	**Caregivers and families** – Caregivers of children between the ages of ten and 24 months who had recently experienced a court-ordered change in primary caregiver. (Birth parents, foster parents and kinship carers were all eligible.) **Sample characteristics** Age – Infant mean age in months (standard deviation) - Control 18.06 (4.49), intervention 17.96 (4.97). Caregiver mean age in years (standard deviation) - Control 36.50 (10.95), intervention 35.39 (10.98). Sex – Male target child – Control n=55 (52.4%), intervention n=63 (60%). Gender of caregivers not reported, only that 5 of the birth parents were fathers. Ethnicity – Ethnicity – Infant - Control - Hispanic n=12 (52.4%), Native American/Alaskan native n=5 (4.8%), Black n=14 (13.3%), mixed race n=18 (17.1%), Native Hawaiian/Other Pacific Islander n=0 (0%), 'unable to determine' n=4 (3.4%), White n=65 (61.9%). Intervention - Hispanic n=9 (8.6%), Native American/Alaskan native n=9 (8.6%),	beneficial effect of the intervention. Based on data from 175 dyads (control n=89, intervention n=86). Sensitivity: Significant difference between groups in favour of the intervention group, small to medium effect size - F=5.22, p=.024, d=0.41. Support: No significant differences between groups - F=0.48, p=.491, d=0.11. Commitment: No significant differences between groups - F=0.86, p=.354, d=–0.17. Understanding of toddlers: Significant difference between groups in favour of the intervention group, small to medium effect size - F=4.21, p=.042, d=0.36. Caregiver outcomes 6 months post-intervention (adjusted for baseline score, age of child, multiple removals, caregiver type, and time between baseline and six months post-intervention assessment). Positive effect sizes represent a beneficial effect of the intervention. Based on data from 129 dyads (control n=70, intervention n=59). Sensitivity: No significant differences between groups - F=2.02, p=.158, d=0.29. Support: No significant differences between groups - F=0.58, p=.446, d=0.18. Commitment: No significant differences between groups - F=0.67, p=.414, d=0.16. Understanding of toddlers: No significant differences between groups - F=3.55, p=.062, d=0.39. Parameter estimates of intervention effects on monthly change – model 1 – estimated effects of intervention on change baseline to immediate post-intervention assessment and immediate post-intervention assessment to 6 months post-intervention assessment. Intervention effect on monthly rate of change - baseline to immediate post-intervention assessment	Key study limitations: Relatively high attrition rate from the study. Relatively high numbers of Participants did not complete follow up assessments (n=28 (26%) at post-intervention stage, and n=34 (32%) at the 6 month point.) At the immediate post-intervention assessment this was comparable by group but at the 6 month point greater numbers of dyads in the intervention group (n=22) failed to complete assessments than those in the control group (n=12). In addition, significant numbers of dyads were excluded from the analyses due to changes in caregiver throughout the course of the study. This meant that data from only 56% of participants in intervention group, and 66% of comparison group were used in the analysis.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Black n=17 (16.2%), mixed race n=23 (21.9%), Native Hawaiian/Other Pacific Islander n=2 (1.9%), 'unable to determine' n=3 (2.9%), White n=51 (48.6%). Religion/belief – Not reported. Disability – Not reported. Long term health condition – Not reported. Sexual orientation – Not reported. Socioeconomic position – Household income of less than $20,000 per year – control n=27 (26.5%), intervention n=23 (23%). Mean caregiver years in education (standard deviation) – control 12.93 (1.79), intervention 13.11 (2.10). Type of abuse – Not reported. Looked after or adopted status – Removed from birth parents home on more than one occasion – control n=5 (4.8%), intervention n=17 (16.2%). chi-square (1, n=210)=7.31, p<.01. Caregiver type – Control – biological parent n=29 (27.6%), kinship carer n=30 (28.6%), foster parent n=46 (43.8%). Intervention - biological parent n=27 (25.7%), kinship carer n=35 (33.3%), foster parent n=43 (41%).	Sensitivity (caregiver): The intervention had a significant positive effect on sensitivity from baseline to post-intervention - b = .34, p<.05. Support (caregiver): The intervention did not have a significant effect on support from baseline to post-intervention - b=.01. Commitment (caregiver): The intervention did not have a significant effect on commitment from baseline to post-intervention - b=.00. Understanding of toddlers (caregiver): The intervention had a significant positive effect on understanding of toddlers from baseline to post-intervention - b=.53, p<.01. Intervention effect on monthly rate of change - immediate post-intervention assessment to 6 months post-intervention assessment Sensitivity (caregiver): The intervention did not have a significant effect on sensitivity from post-intervention to six months - b=-.08. Support (caregiver): The intervention did not have a significant effect on support from post-intervention to six months - b=-.01. Commitment (caregiver): The intervention did not have a significant effect on commitment from post-intervention to 6 months - b=.02. Understanding of toddlers (caregiver): The intervention did not have a significant effect on understanding of toddlers from post-intervention to 6 months - b=-.01. Model 2 – estimated effects of intervention on monthly rate of change - baseline to 6 months post-intervention assessment Sensitivity (caregiver): The intervention did not have a significant effect on sensitivity from baseline to 6 months - b=.06.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Mean age of child in months at first removal (standard deviation) – control 10.86 (7.07), intervention 10.73 (7.78). Mean number of changes in caregivers since birth – Control 2.70 (1.51), intervention 2.67 (1.66). Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size** Comparison numbers – n=105 randomised. NB Only data from dyads which were remained intact were used in analyses. At the immediate post-intervention assessment only 89 dyads remained intact. At the 6 month assessment only 70 remained intact. Intervention number – n=105 randomised. At the immediate post-intervention assessment only 86 dyads remained intact. At the 6 month assessment only 59 remained intact. Sample size – n=210 randomised. NB Only data from dyads which remained intact were used in analyses. At the immediate post-intervention assessment only 175 dyads remained intact. At the six month assessment only 129 remained intact. **Intervention**	Support (caregiver): The intervention did not have a significant effect on support from baseline to 6 months - b=.02. Commitment (caregiver): The intervention did not have a significant effect on commitment from baseline to 6 months - b=.01. Understanding of toddlers (caregiver): The intervention had a marginally significant positive effect on understanding of toddlers from baseline to six months - b=.13, p<.10. NB Using a less commonly accepted significance threshold. **Children and young people's health and wellbeing outcomes** – Child outcomes (adjusted for baseline score, age of child, multiple removals, caregiver type, and time between baseline and post-intervention assessment). Positive effect sizes represent a beneficial effect of the intervention. Based on data from 175 dyads (control n=89, intervention n=86). Security: No significant differences between groups - F=0.68, p=.410, d=0.16. Engagement: No significant differences between groups - F=0.76, p=.386, d=-0.15. Competence: Significant differences between groups in favour of the intervention group, small to medium effect size - F=4.77, p=.031, d=0.42. Problem behaviour: No significant differences between groups - F=0.01, p=.924, d=-0.02. Child outcomes 6 months post-intervention (adjusted for baseline score, age of child, multiple removals, caregiver type, and time between baseline and 6 months post-intervention assessment). Positive effect sizes represent a beneficial effect of the intervention.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Describe intervention – Promoting First Relationships (manualised). **Delivered by** – 'Masters prepared' providers working for 1 of 5 community mental health agencies. Providers were trained for a total of 90 hours over 6 months and were mentored during the provision of the intervention to 3 families. Weekly reflective sessions with other providers were also conducted throughout the programme (p274). **Delivered to** – Caregivers of children between the ages of ten and 24 months who had recently experienced a court-ordered change in primary caregiver. (Sample included birth parents, foster parents and kinship carers.) **Duration, frequency, intensity** – Ten weekly sessions around 60–75 minutes in length. The authors report that 71% of the caregivers assigned to the intervention group participated in all ten sessions, 3% took part in more than half of the sessions, 20% took part in fewer than half of the sessions and 7% did not take part in any sessions. **Key components and objectives of intervention**	Based on data from 129 dyads (control n=70, intervention n=59). Security: No significant differences between groups - F=0.12, p=.736, d=-0.13. Engagement: No significant differences between groups - F=0.71, p=.402, d=–0.18. Competence: No significant differences between groups - F=0.63, p=.429, d=–0.16. Problem behaviour: No significant differences between groups - F=0.62, p=.434, d=–0.16. Internalizing problems: No significant differences between groups - F=0.02, p=.879, d=0.03. Externalising problems: No significant differences between groups - F=0.42, p=.520, d=0.13. Sleep problems: No significant differences between groups - F=2.85, p=.094, d=0.34. 'Other problems': No significant differences between groups - F=0.51, p=.475, d=0.14. Emotional regulation: No significant differences between groups - F=1.02, p=.314, d=0.20. Orientation: No significant differences between groups - F=0.13, p=.723, d=0.06. Parameter estimates of intervention effects on monthly change – model 1 – estimated effects of intervention on change baseline to immediate post-intervention assessment and immediate post-intervention assessment to 6 months post-intervention assessment. Intervention effect on monthly rate of change - baseline to immediate post-intervention assessment – Engagement (child): The intervention did not have a significant effect on engagement from baseline to post-intervention - b=-.02. Competence (child): The intervention did not have a significant effect on competence from baseline to post-intervention - b=.10. Problem behaviour (child): The intervention did not have a significant effect on problem behaviour from baseline to post-intervention - b=.05. Intervention effect on	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	– Although the intervention is manualised the authors report that the sessions were designed specifically for this project. The intervention also includes the use of 5 videotaped child-caregiver interactions which are used to guide sessions and prompt discussion of the strengths of the parent and their interpretation of the child's behaviour. Providers also practised the 'Promoting First Relationships Ways of Being' which emphasises the importance of establishing an emotional connection with the caregiver, sensitive interviewing techniques, reflective practice, positive and instructive feedback, reflection, and responsive and validating statements. **Content/session titles** – The Promoting First Relationships manual includes topics such as attachment theories, challenging behaviours, emotional and social needs of children, and reflective parenting. Topics developed specifically for this study included 'Staying connected during difficult moments' and 'Memory of a strong emotion' (p275). **Location/place of delivery** – Caregiver's home. **Describe comparison intervention**	monthly rate of change - immediate post-intervention assessment to 6 months post-intervention assessment – Security (child): The intervention did not have a significant effect on security from post-intervention to 6 months - b=.01. Engagement (child): The intervention did not have a significant effect on engagement from post-intervention to six months - b=-.01. Competence (child): The intervention had a significant negative effect on competence from post-intervention to 6 months - b=-.15, < .01. Problem behaviour (child): The intervention did not have a significant effect on problem behaviour from post-intervention to 6 months - b=.03. Model 2 – estimated effects of intervention on monthly rate of change - baseline to 6 months post-intervention assessment – Security (child): The intervention did not have a significant effect on security from baseline to 6 months - b=.00. Engagement (child): The intervention did not have a significant effect on engagement from baseline to 6 months - b=-.01. Competence (child): The intervention did not have a significant effect on competence from baseline to 6 months - b=-.07. Problem behaviour (child): The intervention did not have a significant effect on problem behaviour from baseline to 6 months - b=.03. **Caregiver/parent health and wellbeing outcomes** – Caregiver outcomes post-intervention (adjusted for baseline score, age of child, multiple removals, caregiver type, and time between baseline and post-intervention assessment). Positive effect sizes represent a beneficial effect of the intervention. Based on data from 175 dyads (control n=89, intervention n=86). Stress related to perceptions of caring for a difficult child: No significant differences between groups -	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	– Early Education Support. Few details on the comparison intervention are provided. Delivered by – Providers worked at one of five community mental health agencies and had a bachelors degree. Delivered to - Caregivers of children between the ages of ten and 24 months who had recently experienced a court-ordered change in primary caregiver. (Sample included birth parents, foster parents and kinship carers.) **Duration, frequency, intensity** Three monthly sessions of 90 minutes. The authors report that 81% of caregivers in the comparison group participated in all 3 visits, 15% took part in one or two and 4% did not take part in any. **Key components and objectives of intervention** Not reported. The only information provided is that providers of this intervention helped the families to access other services such as Early Head Start, mental health services housing, etc. and suggested activities to '… promote growth and development.' (p 275). **Content/session titles** N/A.	F=1.54, p=.216, d=-0.22. Stress related to perceptions of a dysfunctional caregiver-child interactions: No significant differences between groups - F=.51, p=.478, d=–0.13. Caregiver outcomes 6 months post-intervention (adjusted for baseline score, age of child, multiple removals, caregiver type, and time between baseline and 6 months post-intervention assessment). Positive effect sizes represent a beneficial effect of the intervention. Based on data from 129 dyads (control n=70, intervention n=59). Stress related to perceptions of caring for a difficult child: No significant differences between groups - F=0.07, p=.790, d=0.06. Stress related to perceptions of dysfunctional caregiver-child interactions: No significant differences between groups - F=0.67, p=.415, d=-0.17. Parameter estimates of intervention effects on monthly change – model 1 – estimated effects of intervention on change baseline to immediate post-intervention assessment and immediate post-intervention assessment to 6 months post-intervention assessment. Intervention effect on monthly rate of change - baseline to immediate post-intervention assessment – significant effect on stress related to perceptions of caring for a difficult child from baseline to post-intervention - b=.19. Stress related to perceptions of dysfunctional caregiver-child interactions (caregiver): The intervention did not have a significant effect on stress relayed to perceptions of a dysfunctional caregiver-child relationship from baseline to post-intervention - b=-.05. Intervention effect on monthly rate of change - immediate post-intervention assessment to 6 months post-intervention assessment – Stress related to perceptions of caring for a difficult child (caregiver): The intervention	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Location/place of delivery Caregiver's home. **Outcomes measured** Quality of parenting and parent-child relationships – Sensitivity (e.g. positive interactions, mutuality, verbal and non-verbal support, sensitive instruction) was measured using the Nursing Child Assessment Teaching Scale an observational videotaped assessment measure (Barnard 1994). Support (e.g. acceptance/warmth, follows child lead) was measured (observational) using the Indicator of Parent–Child Interaction (Baggett et al. 2009). Commitment to the child (e.g. missing the child, desire to take care of child in the future) was measured using the caregiver completed This Is My Baby scale (Bates, 1998; Dozier and Lindhiem 2006). Higher scores suggest higher levels of commitment. Understanding of toddlers (emotional and social needs and developmentally appropriate expectations) was measured using the caregiver completed Raising a Baby (Kelly and Korfmacher 2008). Children and young people's health and wellbeing outcomes – Attachment security was measured using the Toddler Attachment Sort-45	did not have a significant effect on stress related to perceptions of caring for a difficult child from post-intervention to 6 months - b=-.08. Stress related to perceptions of dysfunctional caregiver-child interactions (caregiver): The intervention did not have a significant effect on stress related to perceptions of a dysfunctional caregiver-child relationship from post-intervention to 6 months - b=.05. Model 2 – estimated effects of intervention on monthly rate of change - baseline to 6 months post-intervention assessment – Stress related to perceptions of caring for a difficult child (caregiver): The intervention did not have a significant effect on stress related to perceptions of caring for a difficult child from baseline to 6 months - b=.00. Stress related to perceptions of dysfunctional caregiver-child interactions (caregiver): The intervention did not have a significant effect on stress related to perceptions of a dysfunctional caregiver-child relationship from baseline to 6 months - b=.07. **Narrative findings - effectiveness** – Quality of parenting and parent-child relationships Caregiver outcomes post-intervention – Based on data from 175 dyads (control n=89, intervention n=86). At the immediate post-intervention assessment there was a significant difference between groups in scores of caregiver sensitivity (small to medium effect size) measured using the Nursing Child Assessment Teaching Scale and caregiver understanding of toddlers (small to medium effect size) measured using the Raising a Baby scale. There were no significant differences between groups in scores of caregiver support (measured using the Indicator of Parent–Child Interaction) and caregiver commitment to the child (measured	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	(Kirkland et al. 2004), a modified version of the Attachment Q-Sort (Waters 1987). The authors report the use of a sorting technique described by the developers of TAS45 as trilemmas. These are a specific set of three from which the observer chooses which is most like and which is least like the behaviour of the child. Each of descriptive statement appears in two trilemmas and there are 30 trilemmas in total. Scoring provides an overall score of attachment security. Engagement (e.g. positive feedback, turn-taking, etc.) was measured (observation) using scores from the Indicator of Parent–Child Interaction (Baggett et al. 2009) as used to measure caregiver support. Competence (e.g. positive social behaviours) and Problem-Behaviours in the last month were measured using the Brief Infant Toddler Social and Emotional Assessment (Briggs-Gowan and Carter 2002). Child behaviour in the last two months was measured using the caregiver completed the Child Behavior Checklist for Ages 1 ½–5 (Achenbach and Rescorla 2000). The internalizing, externalizing, sleep problems and 'other problems' scales were used. NB This was only measured at the 6 months post-intervention assessment point due to the	using the This Is My Baby scale). Caregiver outcomes six months post-intervention - Based on data from 129 dyads (control n=70, intervention n=59). At the 6 months post-intervention assessment there were no significant differences between groups in scores of caregiver sensitivity, caregiver support, caregiver commitment, and caregiver understanding of toddlers. Caregiver/parent health and wellbeing Caregiver outcomes post-intervention – Based on data from 175 dyads (control n=89, intervention n=86). At the immediate post-intervention assessment there were no significant differences between groups in scores of caregiver stress related to perceptions of caring for a difficult child or caregiver stress related to perceptions of a dysfunctional caregiver-child relationship (both measured using the short form of the Parenting Stress Index). Caregiver outcomes 6 months post-intervention - Based on data from 129 dyads (control n=70, intervention n=59). At the 6 months post-intervention assessment there were no significant differences between groups in scores of caregiver stress related to perceptions of caring for a difficult child or caregiver stress related to perceptions of a dysfunctional caregiver-child relationship. Children and young people's health and wellbeing Child outcomes post-intervention - based on data from 175 dyads (control n=89, intervention n=86). At the immediate post-intervention assessment there was a significant difference between groups in scores of child competence in favour of the intervention group (small to medium effect size) measured using the Brief Infant Toddler Social and Emotional Assessment tool. There were no significant differences between groups in scores of child attachment security (measured using the Toddler Attachment Sort-45), child engagement	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	age of the child at baseline and immediate post-intervention assessment. Emotional regulation was measured using seven out of ten items from the Emotional Regulation factor of the Bayley-III Screening Test (Bayley 2005). NB This was not measured at the immediate post-intervention assessment as the authors report the interval was too brief. Orientation/engagement was measured using 6 out of 9 items from the Orientation/Engagement factor of the Bayley Behavior Rating Scales (Bayley 1993). NB This was not measured at the immediate post-intervention assessment as the authors report the interval was too brief. Caregiver/parent health and wellbeing outcomes – Parenting stress linked to the perception of caring for a difficult child or a dysfunctional caregiver–child relationship was measured using the caregiver completed short form of the Parenting Stress Index (Abidin 1995) administered during the structured interview.	(measured using the Indicator of Parent–Child Interaction) or child problem behaviours (measured using the Brief Infant Toddler Social and Emotional Assessment). Child outcomes six months post-intervention - Based on data from 129 dyads (control n=70, intervention n=59). At the 6 months post-intervention assessment there were no significant differences between groups in scores of child attachment security, child competence or child problem behaviours. At the 6 months post-intervention assessment the study also measured internalizing problems externalizing problems, sleep problems, 'Other problems' (all measured using subscales of the Child Behavior Checklist for Ages 1 ½–5) and found no significant differences between groups in scores on any of these scales. Emotional regulation and orientation were also measured (both using the Orientation/Engagement factor of the Bayley Behavior Rating Scales) and no significant differences between groups in scores on either of these were found. Parameter estimates of intervention effects on monthly change – model 1 – estimated effects of intervention on change baseline to immediate post-intervention assessment and immediate post-intervention assessment to 6 months post-intervention assessment. Intervention effect on monthly rate of change - baseline to immediate post-intervention assessment – Between baseline assessments and immediate post-intervention assessments the intervention had a significant positive effect on monthly change in scores of caregiver sensitivity and caregiver understanding of toddlers. The intervention did not have a significant effect on monthly change in scores of caregiver support, caregiver commitment, caregiver stress related to perceptions of caring for a difficult child, caregiver stress	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		related to perceptions of a dysfunctional caregiver-child relationship, child attachment security, child engagement, child competence, and child problem behaviours. Intervention effect on monthly rate of change - immediate post-intervention assessment to 6 months post-intervention assessment – Between immediate post-intervention assessments and 6 months post-intervention assessments the intervention had a significant negative effect on monthly change in scores of child competence. The intervention did not have a significant effect on monthly change in scores of caregiver sensitivity, caregiver support, caregiver commitment, caregiver understanding of toddlers, caregiver stress related to perceptions of caring for a difficult child, caregiver stress related to perceptions of dysfunctional caregiver-child interactions, child attachment security, child engagement, and child problem behaviours. Model 2 – estimated effects of intervention on monthly rate of change - baseline to six months post-intervention assessment – Between baseline assessments and 6 months post-intervention assessments the intervention had a marginally significant positive effect on monthly change in scores of caregiver understanding of toddlers (using a less commonly accepted significance threshold of $p < .10$). The intervention did not have a significant effect on monthly change in scores of caregiver sensitivity, caregiver support, caregiver commitment, caregiver stress related to perceptions of caring for a difficult child, caregiver stress related to perceptions of dysfunctional caregiver-child relationships, child security, child engagement, child competence, and child problem behaviours.	

19b. Spieker SJ, Oxford ML, Fleming CB (2014) Permanency outcomes for toddlers in child welfare two years after a randomized trial of a parenting intervention. Children and Youth Services Review, 44: 201–6

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim Study aim: 'We ask whether an attachment-based parenting program directed at promoting sensitive caregiving has an effect on the child's placement stability and legal permanency two years later' (p202). **Methodology** RCT. **Country** Not United Kingdom. USA. **Source of funding** Government.	**Participants** Caregivers and families. The intervention aims to increase parent sensitivity and to see whether the intervention has an effect on the permanency of foster placements. **Sample characteristics** Age - Child Intervention Mean age - 17.96 months Control Mean age- 18.06 months Caregiver Intervention Mean age - 35.39 years Control Mean age- 36.50 years Sex - Intervention - 63 male Control - 55 male 105 in each group Ethnicity - Intervention Native American/Alaskan native 9 Black 17 Mixed race 23 Native Hawaiian/Other Pacific islander 2 Unable to determine 3 White 51 Control Native American/Alaskan native 5 Black 14 Mixed race 18 Native Hawaiian/Other Pacific islander 0 Unable to determine 4 White 65 Looked after or adopted status - All the children in the study were in foster care. **Sample size** Comparison numbers - 105 children	**Effect sizes** Quality of parenting and parent-child relationships Logistic regression intent-to-treat models were constructed to assess differences by intervention in stability and permanency. Model 1 entered intervention and all covariates (foster/kin placement, age of child in months, months in child welfare, number of prior placements, multiple removals and commitment) as predictors. Model 2 entered an additional interaction term between caregiver type and intervention condition. 1. Stability Model 1 found no significant effect on stability of the intervention. Significant effects were found in two covariate areas: Children who enrolled with birth parents were found to be more likely to be living with them two years later (OR=0.16, 95% CI 0.07 to 0.34). Also, those children living with caregivers who had a greater 'commitment level' were also more likely remain living with them 2 years later (OR=1.89, 95% CI 1.20 to 2.98). Model 2 suggested a differential relationship with stability depending on which type of carer received the intervention, with the intervention x caregiver interaction term showing a marginally statistically significant result (p<0.10 OR=3.91, 95% CI 0.82 to 18.52). This showed that being in the intervention group was more positively related to placement stability for foster/kin caregivers compared to birth parents. 2. Permanency Model 1 showed no impact of intervention group on permanency. In Model 2 the intervention x caregiver interaction term was significant (OR=9.67, 95% CI 1.54 to 60.68), suggesting that being in the intervention group was more positively related to permanency for foster/kin carers compared to birth carers. 3.	**Overall assessment of internal validity** - **Overall assessment of external validity** ++ **Overall validity score** - Coded for consistency with Spieker (2012).

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Intervention numbers - 105 children **Intervention** Describe intervention The intervention was called Promoting First Relationships (PFR). It is designed to improve the sensitivity and responsiveness of caregivers toward toddlers that have been placed in their care. The intervention is made up of 10 sessions (over 10 weeks). The sessions lasted 60 to 75 minutes. The training was delivered by a mental health worker. The intervention uses: 'attachment theory-informed, strength-based consultation strategies in conjunction with video feedback' (p203). Five sessions used 'reflective video feedback' (p202). The videos of caregivers playing with the child were watched by the caregivers and the mental health worker. There was discussion on parenting strengths and how well the care giver responded to the child's cues. The caregivers reflected on the previous session's progress. Caregivers were issues with up to 15 handouts on various parenting topics. PFR aimed to inform caregivers that difficult or challenging behaviour from toddlers was a sign of attachment	Understanding the caregiver x intervention group interaction Follow-up logic regression was also carried out on all cases involving foster/kin carers. Here a significant effect of intervention was found in relation to permanency (OR=3.83, 95% CI 1.07 to 13.78). Carers' commitment also produced an effect. There was a tendency for foster/kin carers in the PFR group to report higher levels of commitment than those in the EES group (F(1, 82)=2.90, p=0.09).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	needs or a 'language of distress', re-curring greater emotional openness from the caregiver. **Delivered by** Mental health practitioners, at masters level. **Delivered to** Caregivers of toddlers who had been placed with them by child welfare. Duration, frequency, intensity, etc. 10 sessions over 10 weeks. **Location/place of delivery** In the family home. Describe comparison intervention The comparison group received Early Education Support (EES). Provided from a community agency. Three sessions, over three 3, in the family home. Delivered by a child development specialist. The specialist assisted with suggesting activities that might stimulate the child's development and signposted to services in the community. **Outcomes measured** Quality of parenting and parent-child relationships The study sought to measure stability and permanency outcomes. Stability		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	was defined as whether the child had remained with the same caregiver since randomisation in to the study. Permanency included reunification and discharge to the study birth parent, adoption by the study kin or non-kin caregiver or legal guardianship by the study caregiver.		

20. Stronach EP, Toth SL, Rogosch F et al. (2013) Preventive interventions and sustained attachment security in maltreated children. Development and Psychopathology 25: 919–30

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim – The aim of the study was to evaluate the efficacy of a child–parent psychotherapy programme and a psychoeducational parenting intervention in comparison to care as usual. The study focused on secure attachment and behavioural functioning at 12 months (expanding on the findings from a previous study in which as-	**Participants** **Children and young people** – Maltreated children or siblings of maltreated children living with their biological mother (service records were used to identify families in which an indicated report for abuse or neglect had been filed) were randomised to one of three treatment conditions. The study also included a group of non-maltreated children in receipt of Temporary Assistance to Needy Families '... in order to obtain a demographically similar comparison group' (p 922). Service records for the Temporary Assistance to Needy Families programme were screened to exclude	**Effect sizes** **Quality of parenting and parent-child relationships** – NB It is not possible to provide a narrative description of the effect size due to lack of detail on the type of effect size reported. Children's attachment in the 'Strange Situation' assessment at 12 months post-intervention follow-up (treatment completer analysis) – Differences in attachment classification by group – Significant difference, chi-square (9, n=145) = 33.49, p<.001, ES=0.28. Differences in rates of secure attachment between Child-parent Psychotherapy group and care as usual group – Children in the Child-parent Psychotherapy group had significantly higher rates of secure attachment than those in the care as usual group, chi-square (1, n=76) = 16.33, p<.001, ES=0.46.	**Overall assessment of internal validity** + **Overall assessment of external validity** ++ **Overall validity score** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
sessments were conducted in the immediate post-intervention period - Cicchetti D, Rogosch F, Toth SL, 2006, Fostering secure attachment in infants in maltreating families through preventive interventions. Development and Psychopathology 18: 623–49). NB The NCCSC have reported the findings of this paper as part of the findings extracted from Goldman Fraser J, Lloyd SW, Murphy RA et al. (2013) Child exposure to trauma: Comparative effectiveness of interventions addressing maltreatment. Comparative Effectiveness Review 89: 1–161. **Methodology** RCT. – Maltreated children and their mothers were randomised to either a child-parent psychotherapy programme, a psy-	families in which there was a history of maltreatment. **Caregivers and families** – Biological mothers of maltreated children (residing together) they or siblings of maltreated children living with their biological mother (service records were used to identify families in which an indicated report for abuse or neglect had been filed) were randomised to one of three treatment conditions. The study also included a group of non-maltreated children and their mothers in families in receipt of Temporary Assistance to Needy Families '...in order to obtain a demographically similar comparison group' (p922). **Sample characteristics** Age – All groups - Age of child at baseline assessment (approximate) - 13.31 months, SD=0.81. Age of mother at baseline assessment (approximate) - 26.98 years, SD=5.98. Sex – Maltreated group of children - female n=77, male n=60. Non-maltreated group of children - female n=24, male n=28. All caregivers were female. Ethnicity – All groups - Mothers - 74% were from a 'minority race'. Children - not	Differences in rates of disorganised attachment between Child-parent Psychotherapy group and care as usual group - Children in the Child-parent Psychotherapy group had significantly lower rates of disorganised attachment than those in the care as usual group, chi-square (1, n=76) = 3.83, p=.05, ES=0.23. Differences in rates of secure attachment between Psychoeducational Parenting Intervention group and care as usual group – Non-significant difference, chi-square (1, n=71) = 1.27, p=.26, ES = 0.13. Differences in rates of disorganised attachment between psychoeducational parenting intervention group and care as usual group - Non-significant difference, chi-square (1, n=71) = 0.62, p=.43, ES=0.09. Differences in rates of secure attachment between Child-parent Psychotherapy group and psychoeducational parenting intervention group - Children in the Child-parent Psychotherapy group had significantly higher rates of secure attachment than those in the Psychoeducational Parenting Intervention group, chi-square (1, n=49)=5.41, p=.02, ES=0.33. Differences in rates of disorganised attachment between child-parent psychotherapy group and psychoeducational parenting intervention group - Children in the Child-parent Psychotherapy group had significantly lower rates of disorganised attachment than those in the Psychoeducational Parenting Intervention group, chi-square (1, n=49) = 5.52, p=.02, ES=0.34. Differences in rates of secure attachment between child-parent psychotherapy group and non-maltreated comparison group - Non-significant difference, chi-square (1, n=74) = 2.06, p=.15, ES=0.17. Differences in rates of secure attachment between psychoeducational parenting intervention group and non-maltreated comparison group - Non-significant difference, chi-square (1,	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
choeducational parenting intervention or care as usual. The study also included a group of non-maltreated children and their mothers as a comparison. **Country** Not UK. – USA. **Source of funding** Government. – Administration for Children, Youth and Families and the National Institute for Mental Health. Other. – Spunk Fund, Inc.	reported. Religion/belief – Not reported. *Disability* – Not reported. Long term health condition – Long term health condition - All groups - Mothers with a lifetime history of post-traumatic stress disorder - 35% (n=65). Mothers meeting criteria for post-traumatic stress disorder at baseline assessment - 20% (n=38) Sexual orientation – Not reported. Socioeconomic position – Socioeconomic position - All groups - Mothers who were married - 12.7%. Mothers with a high school diploma or equivalent - 58.2%. Families in receipt of Temporary Assistance to Needy Families - 96.3%. Type of abuse – Maltreated group - Participating child indicated as target of abuse and/or neglect in family report - 66.4%. Sibling of participating child indicated as target of abuse and/or neglect in family report 33.6%. Participating child who had experienced neglect - 84.6%. Participating child who had experienced emotional maltreatment - 69.2%. Participating child who had experienced physical abuse - 8.8%. Participating child who had experienced sexual	n=69) = 1.64, p=.20, ES=0.13. Differences in rates of secure attachment between care as usual group and non-maltreated comparison group – Children in the non-maltreated comparison group had significantly higher rates of secure attachment than those in the care as usual group, chi-square (1, n=96) =8.68, p=.003, ES=0.30. Differences in rates of disorganised attachment between child-parent psychotherapy group and non-maltreated comparison group – Non-significant differences, chi-square (1, n=74), = 0.13, p= .72, ES=0.04. Differences in rates of disorganised attachment between psychoeducational parenting intervention group and non-maltreated comparison group – Children in the Psychoeducational Parenting Intervention group had significantly higher rates of disorganised attachment than those in the non-maltreated comparison group, chi-square (1, n=69) = 5.40, p=.02, ES=0.28. **Children and young people's health and wellbeing outcomes** – Child behaviour problems (maternal perceptions) at 12 months post-intervention follow-up (treatment completer analysis) – Multivariate effect for treatment group – Non-significant effect, Wilks λ=0.95, F (9, 145)=0.86, p=.56. Child behaviour problems (maternal perceptions) at 12 months post-intervention follow-up (intent-to-treat analysis) - Multivariate effect for treatment group - Wilks λ=0.93, F (9, 145)=1.15, p=.32. NB. Intent-to-treat analysis does not include all comparisons used in treatment completer analysis.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	abuse - 0%. All groups - Mothers who had been maltreated as children - 79.4% (n=150). Looked after or adopted status – All groups - All children were living with their biological mother. Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size** Comparison numbers – Care as usual group (community standard) - n=35. Non-maltreated comparison group - n=52. Intervention number – Child–parent psychotherapy - n=53. Psychoeducational parenting intervention - n=49. Sample size – Maltreated sample - n=137. Non-maltreated sample - n=52. Total sample - n=189. **Intervention category** Parenting programmes – The study evaluated 2 interventions - Child–parent psychotherapy and a psychoeducational parenting intervention. Child-Parent Psychotherapy – The study evaluated 2 interventions - Child–parent psychotherapy and a	**Narrative findings - effectiveness** – Children's attachment in the 'Strange Situation' assessment at 12 months post-intervention follow-up (treatment completer analysis) At 12 month post-intervention follow-up there were significant differences between groups in attachment classifications. Maltreated children randomised to the child-parent psychotherapy group showed significantly higher rates of secure attachment and significantly lower rates of disorganised attachment than those assigned to the care as usual group. There was no significant differences in rates of secure or disorganised attachment between maltreated children randomised to the psychoeducational parenting intervention group or those randomised to the care as usual group. Maltreated children randomised to the child-parent psychotherapy group showed significantly higher rates of secure attachment and significantly lower rates of disorganised attachment than those assigned to the psychoeducational parenting intervention group. There were no significant differences in rates of secure attachment between maltreated children randomised to either the child-parent psychotherapy group or the psychoeducational parenting intervention group in comparison to the non-maltreated comparison group. Children in the non-maltreated comparison group had significantly higher rates of secure attachment than maltreated children randomised to the care as usual group. There were no significant differences in rates of disorganised attachment between maltreated children randomised to the Child-Parent Psychotherapy group and those in the non-maltreated comparison group. Maltreated children randomised to the Psychoeduca-	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	psychoeducational parenting intervention. **Intervention** **Describe intervention** – Child–parent psychotherapy (manualised) - A 'therapeutic model' which is intended to '... enrich the complex relationship between traumatized children and parents' (p923). Psychoeducational parenting intervention (manualised) - Modelled after a nurse home visiting programme (Olds DL and Kitzman H, 1990, Can home visitation improve the health of women and children at environmental risk? Pediatrics 86: 108–16) which provided low-income mothers with education about the physical and psychological development of their infant. **Delivered by** – Child–parent psychotherapy - Therapists (masters level). Psychoeducational parenting intervention - Therapists (masters level). **Delivered to** – Child–parent psychotherapy - Mother and child. Psychoeducational parenting intervention - Mothers. **Duration, frequency, intensity** – Child–parent psychotherapy - weekly	tional Parenting Intervention group showed significantly higher rates of disorganised attachment in comparison to children in the non-maltreated comparison group. Child Behaviour problems at 12 months post-intervention follow-up (treatment completer analysis) At 12 month post-intervention follow-up there were no significant differences between groups in maternal perceptions of externalising, internalising or total child behaviour problems.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	sessions delivered over a 12-month period (no further details provided). The authors report that n=32 (60.4%) of dyads randomised to the Child–parent psychotherapy took part in the programme. Psychoeducational parenting intervention - weekly sessions delivered over a 12-month period (no further details provided). The authors report that n=24 (48.9%) of dyads randomised to the Child–parent psychotherapy took part in the programme. **Key components and objectives of intervention** – Child–parent psychotherapy - The therapy provided is 'non-directive' and 'non-didactic' and focuses on the mother's negative perceptions of her relationship with her child stemming from her own negative experiences or insecure representational model. During the sessions the therapist observes interactions between the mother and child (using the child's own toys) and responds 'empathically' to these. Mothers are able to use this support to address their negative perceptions and differentiate '… between affect and impulses associated with past experiences versus current relationships …' (p923) which results in more positive and sensitive interac-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	tions. The authors also note that further individual sessions with mothers could be arranged to '… discuss content that may be distressing to a verbal child, such as the mother's physical or sexual abuse history' (p923). Psychoeducational parenting intervention - The psychoeducational parenting intervention is didactic and intended to provide mothers with information on child development and parenting techniques, as well as helping them to reduce stress in the parenting role and to increase their personal satisfaction. Therapists also use a range of cognitive and behavioural techniques to help parents improve their parenting and problem solving skills, and to develop methods of relaxation and social support resources. This work was tailored to the primary needs of mothers. **Content/session titles** – N/A. **Location/place of delivery** – Child–parent psychotherapy - Family home. Psychoeducational parenting intervention - Family home. **Describe comparison intervention** – Care as usual - Cases of families randomised to the 'community standard' group were managed by the De-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	partment of Human Services '... according to their customary approach' (p922). They were also helped to access services and support that may have been difficult to reach otherwise. The authors note that this varied from '... minimal contact to group parent skills training or individual counseling' (p922). **Outcomes measured** Quality of parenting and parent-child relationships – Children's attachment was measured during participation in the 'Strange Situation' assessment. These were videotaped and attachment levels was classified into 1 of 5 categories using the Attachment Organization in Preschool Children: Procedures and Coding Manual (Cassidy J and Marvin RS, 1992, Attachment organization in preschool children: Procedures and coding manual. Pennsylvania State University: Unpublished). Children and young people's health and wellbeing outcomes – Maternal perceptions of children's behavioural problems were measured using the Child Behavior Checklist/2-3 (Achenbach 1992). The authors note that they interpreted scores on this scale as maternal perceptions of be-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	havioural problems rather than objective observations because of the tendency for abusive parents to over-report externalising behaviours.		

21. Swenson CC, Schaeffer CM, Henggeler SW et al. (2010) Multisystemic therapy for child abuse and neglect: A randomized effectiveness trial. Journal of Family Psychology 24: 497–507

Study aim To evaluate an adaptation of multisystemic therapy for physically abused adolescents and their families. **Methodology** RCT. **Country** Not UK. – US. **Source of funding** Government. – US National Institute of Mental Health Grant.	**Participants** **Children and young people** – 86 youth (10 to 17 years old) and parent who was implicated in the CPS (Child Protection Services) report of physical abuse. Cases were referred by the county CPS. **Caregivers and families** – Parents. **Sample characteristics** Age – Mean age of youth: 13.88 years (SD 2.07 years); mean age of parents: 41.79 years (SD 10.49 years). Sex – Youths: 55.8% female; parents: parents: 65.1% female, and 58.1% were single parents. Ethnicity – Youths: 68.6% Black, 22.1% White, and 9.3% other; Religion/belief – Not reported Disability	**Effect sizes** **Incidence of abuse and neglect** – Maltreatment outcomes at 16 months reabuse of the youth: No significant difference - 4.5% youth in the MST-CAN condition experienced an incident vs. 11.9% of youth in the EOT condition. However, this was not significant (Chi-square=1.56, p=0.198, ns) odds ratio not reported Reabuse by parents: No significant difference -parents who received EOT were more likely to have an incident of reabuse relative to MST-CAN parents (4.8% vs.2.3% respectively, 95% CI 0.19 –24.43, ns), odds ratio not reported. Youth who received MST-CAN were significantly less likely to experience an out-of-home placement over 16 months than were youth in the EOT condition. (Chi-square=3.74, p<0.05, phi=0.21). **Quality of parenting and parent-child relationships** – Parenting behaviours at 16 months 1. Youth and parent-reported neglect Significantly greater decrease in MSTCAN than EOT group (youth-reported d=0.89; parent-reported d=0.28) 2. Youth-reported psychological aggression Significantly greater decrease in MSTCAN than EOT group (youth-reported d=0.21) 3. Youth-reported minor assault Significantly greater decrease in MSTCAN than EOT group (youth-reported	**Overall assessment of external validity** + **Overall assessment of internal validity** + **Overall validity score** +

	– Not reported Long term health condition – 28% of youth received medication for attention deficit hyperactivity disorder [ADHD] and 7% of caregivers for depression or anxiety. Sexual orientation – Not reported Socioeconomic position – Caregiver: high school graduate 75.0 - 64.3%. Family annual income (US$): < 10,000- range from 19.2% to 31.3%; 10,001–15,000, range from 14.4% to 13.1%; 15,001–20,000, range from 2.4% to 5.2%; 20,001–25,000, range from 19.1% to 7.8%; 25,001–30,000, range from 19.2% to 5.3%; > 30,000, range from 26.4% to 36.5%. Type of abuse – More than 80% of the abuse incidents included at least minor injuries, and 23.3% of families had a prior CPS report. Abuse categories: Pushing or shaking (no injury), Excessive spanking (no injury), Pinched or bit (minor injury), Hit with object (minor injury), Threatened with a weapon, Major assault (e.g., battery, beating). Looked after or adopted status – Some youths on placement Unaccompanied asylum seeking, refugee or trafficked children – No. **Sample size** Comparison numbers – 42 youths	d=0.14) 4. Youth and parent-reported severe assault Significantly greater decrease in MSTCAN than EOT group (youth-reported d=0.54; parent-reported d=0.57). Clinical significant of this: rates of youth-reported incidence of severe assault by parents in the MST-CAN group across the 16 months was approximately half that of the comparison group (4.7 vs 9.8 cases respectively). 5. Youth and parent-reported non-violent discipline Significantly greater decrease in MSTCAN than EOT group (youth-reported d=0.20; parent-reported d=0.57). **Children and young people's health and wellbeing outcomes** – Youth outcomes at 16 months 1. youth-reported PTSD symptoms Significant decrease in MSTCAN group from 17.8% at baseline to 8.9% at 16 months, compared with an increase in EOT group from 19% at baseline to 21.4% at 16 months. Improvement was significantly greater for MST-CAN youth than EOT youth (effect size (d) 0.68 [no CI provided]) 2. youth-reported dissociative symptoms Significant decrease in MSTCAN but not EOT youth (d=0.73) 3. parent-reported internalising Significant decrease in MSTCAN but not EOT youth (d=0.71) 4. parent-reported PTSD Significant decrease in MSTCAN but not EOT youth (d=0.55) 5. parent-reported total symptoms Significant decrease in MSTCAN but not EOT youth (d=0.85) **Caregiver/parent health and wellbeing outcomes** – Parent outcomes at 16 months 1. distress Parental psychiatric distress: decrease in MSTCAN group from 20.5% at baseline to 5.3% at 16 months, compared with little change in EOT group from 16.7% at baseline to 15.8% at 16 months. This amounted to a significantly greater decrease in MSTCAN than EOT parents	

	Intervention number – 44 youths Sample size – 86 youths (44 MST; 42 EOT) **Intervention category** Other – Multisystemic Therapy for Child Abuse and Neglect (MST-CAN) **Intervention** **Describe intervention** – Adapted Multisystemic Therapy for Child Abuse and Neglect (MST-CAN) (Swenson et al. 2010). MST-CAN is an adaptation of Multi-systemic therapy (MST, Henggeler et al. 2009). The principal features of MST include addressing the multidetermined nature of serious clinical problems, working with the family to achieve behaviour changed, delivering services in the home to overcome barriers to service access, integrating evidence-based interventions within the delivery of MST, and using a comprehensive quality assurance system to support therapist fidelity. **Delivered by** – The MST-CAN therapists worked in a team of 3 with full time MST supervisors. An MST-trained psychiatrist was available to the team and provided evidence-based pharmacotherapy to children and parents when warranted. To maintain program fidelity, all MST-	(d=0.63) 2. Social support for parents significant increases reported in MSTCAN parents in total (d=0.46), appraisal (0.67), and belonging social support (0.57), whereas EOT counterparts did not. **Narrative findings - effectiveness** – This RCT reported improved outcomes in youths and parents who were implicated in CPS report of physical abuse, when they were offered Multisystemic therapy for child abuse and neglect (MSTCAN) compared with Enhanced Outpatient Treatment (EOT). At 16 months, MSTCAN was found to be more effective in decreasing youth and parent mental health problems such as internalizing, PTSD and dissociative symptoms. Parental use of severe assault was much reduced (reported by both youth and parents), so was youth-reported reduction in parental neglect after the intervention when compared with EOT. There was also improvement in perceived social support and belonging among MSTCAN parents, not evident in the EOT parents. Fewer youth in MSTCAN group experienced reabuse than those in the EOT group, though the effect was not statistically significant.	

	CAN therapists received a 5-day orientation to the standard MST model. Additional training sessions were provided for the MST-CAN adaptations. Therapists participated in 4 hours of weekly group supervision (two sessions of 2 hours each) and individual supervision as needed. All therapists in both intervention and comparison groups had masters degrees in clinical counselling, social work, or psychology and at least 1 year of prior clinical experience. **Duration, frequency, intensity** – As MST is outcome driven rather than time driven, the treatment length can vary by family. In this study, the length of treatment was allowed to extend beyond the typical 4 to 6 months used in standard MST. The MST therapists delivered interventions in the home and other community locations (e.g., school) at times convenient to families (e.g., evenings, weekend hours). The frequency of treatment sessions was adjusted to family need—ranging from daily sessions to once or twice per week. The team also provided a 24 hour/7 day per week on call service for families to manage crises. **Key components and objectives of intervention** – MST practices aimed at overcoming barriers to service access. MSTCAN		

	used a recursive analytical process to identify, develop, and prioritise interventions. Some of these interventions are conducted with all families and others only as warranted. 1. A safety plan was developed for and agreed by each family that outlined what family members would do if they felt unsafe. 2. The treatment team worked closely with CPS, aiming to foster positive CPS-family relations and ensure that any decisions made by CPS were based on clinical progress made by Support for the parent to address cognitions about the abuse incident, accept responsibility for the abuse, and apologise to the child and family. 4. A number of cognitive behavioral and behavioral interventions incorporated as needed such as CBT for deficits in anger management. Similarly, a CBT protocol was used with families who had problem solving skills or difficulties communicating. In addition, parents experiencing PTSD symptoms received prolonged exposure therapy. **Location/place of delivery** – Community setting (mental health centre), or homes or school as convenient to families. **Describe comparison intervention** – Enhanced Outpatient Treatment (EOT) A standard services provided for physically abused youths and their		

	parents as well as enhanced engagement and parent training interventions. Additional services were provided as needed Referrals made to the Centre were followed up with psychiatric assessment. To support participation and retention in interventions within the EOT condition, therapists made multiple efforts to remind families about upcoming appointments and to reschedule missed appointments. Therapists also made home visits and provided vouchers to cover family transportation to the Centre. Parent training. The Systematic Training for Effective Parenting of Teens (STEP-TEEN; Dinkmeyer et al. 1998) programme was provided for all parents. STEP-TEEN is a structured, 7-lesson or longer, group-based parent-training programme that targets parent-child relations, through didactic instruction, role-play, videotapes and group discussion, the program teaches skills in understanding teens, communication, problem solving, building responsibility, and encouraging cooperation. Within the EOT condition, STEP-TEEN, enhanced engagement, and other standard services delivered at the Centre were provided by a single therapist. All therapists in both intervention and comparison groups had masters degrees in clinical counselling, social work, or psychology and at least 1 year of prior clinical experience. To maintain programme fidelity, therapists		

	in the EOT condition received one day of training on administering the STEP-TEEN programme and participated in weekly 1.5-hour consultation sessions with a supervisor not involved in any clinical aspect of MST-CAN. **Outcomes measured** Incidence of abuse and neglect – Two measures of reabuse: new report of abuse of the target child; abuse of any child by the target parent. Obtained from CPS records. Quality of parenting and parent-child relationships – Parent functioning using the Global Severity Index (GSI) of the Brief Symptom Inventory (BSI; Derogatis 1975), also the number of symptoms was measured on the BSI Positive Symptom Total Scale (PST). Children and young people's health and wellbeing outcomes – Youth functioning, measured using (a) the 113-item Child Behavior Checklist (CBCL; Achenbach 1991) and a 20-item CBCL-PTSD scale (Ruggiero & McLeer 2000); (b) The 54-item Trauma Symptom Checklist for Children (TSCC; Briere 1989) was used to assess children's self-reports of trauma-related symptoms including Anger, Anxiety, Depression, Dissociation, and Posttraumatic Stress; (c) Parent ratings of youth social skills with the Social Skills Rating System (Gresham & Elliott 1990).		

	Caregiver/parent health and wellbeing outcomes – Parenting behaviour (parent self-report and youth report) using the Conflict Tactics Scale (CTS; Straus et al. 1998) including neglect, psychological aggression, minor assault, severe assault, and nonviolent discipline. Social support and belonging as measured by the Interpersonal Support Evaluation List (ISEL; Cohen et al. 1985).		

22. Toth SL, Sturge-Apple ML, Rogosch FA et al. (2015) Mechanisms of change: Testing how preventative interventions impact psychological and physiological stress functioning in mothers in neglectful families. Development and psychopathology 27: 1661–74

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To identify the impact of 2 preventative interventions - Child-Parent psychotherapy (CPP) and Pyschoeducational Parenting Intervention (PPI) on levels of maternal stress in mothers from neglectful families, compared with community standard treatment for maltreating parents (CS) and a nonmaltreating comparison group (NC).	**Participants** Children and young people - Infants. Caregivers and families - Mothers. A Department of Human Services (DHS) recruitment liaison identified infants who were currently residing with biological families who were subject to CPS reports for maltreatment, specifically for this study, neglect. Eligible mothers were contacted and gave written consent for their names to be passed to project staff. A demographically comparable sample of low-income (receiving Temporary Assistance to Needy Families TANF)	**Effect sizes** **Caregiver/parent health and wellbeing outcomes** 1. At baseline: a) Parent-related psychological stress Mothers in CPP group reported higher levels of parent-related psychological stress and child-related stress compared to mothers in NC group (d=2.07). b) Child-related psychological stress CPP mothers also reported higher levels of child-related psychological stress when compared to the NC group and CS groups (d=3.60 and 2.11 respectively). No other differences between groups was noted. 2. Post intervention: a) Parent-related psychological stress	**Overall assessment of internal validity** + More precision on results in providing confidence intervals for differences between groups and a power calculation combined with no blinding. **Overall assessment of external validity** + Lack of UK focus and doesn't exactly address review question.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Methodology: RCT. This RCT randomised 105 mothers known to a US Department of Human Services as their children were subject to CPS reports for maltreatment to either Child-Parent Psychology intervention (CPP) or Psychoeducational Parenting (PPI) or community standard treatment (CS). Total 52 nonmaltreating mothers, enrolled in the Temporary Assistance to Needy Families (TANF) programme were recruited as a comparison group after being screened for child maltreatment through review of DHS record and interviews. The Maltreatment Classification System (Barnett et al. 1993) was used to code DHS records in order to recruit participants. Data was collected pre-intervention (baseline (child mean age 13.3 months));	mother-infant dyads, were recruited after being screened for child maltreatment via DHS and preventive records and a Maternal Maltreatment Classification Interview by DHS and research staff. **Sample characteristics** • Age - Mean age of infants was 13 months - CPP group 13.33; PPI group 13.32; CS group 13.31; non-maltreated comparison group 13.31. Mean maternal age - CPP group 26.98; PPI group 26.35; CS group 27.7; NC group 26.06. • Sex - Infants - 46 boys, 59 girls • Ethnicity - not reported, only that statistical analysis found no difference between the 2 groups. • Religion/belief– Not reported. • Disability– Not reported. • Long term health condition– Not reported. • Sexual orientation– Not reported. • Socioeconomic position– Maternal education attainment at less than High School: CPP group 62.8%; PPI group 44.1%; CS group 51.9%; NC group 26.9% Total income in $1000: CPP group 17.20; PPI group 15.90; CS group 18.54; NC group 16.96.	Mothers in the PPI group reported significantly improved reduction of levels of parenting-related psychological stress, both and pre and post intervention and when compared with the NC and CS groups (whose levels of parenting-stress remained the same or increased respectively)- d=2.43 and 2.44 respectively. There was no significant difference between PPI and CPP mothers (who reported no significant changes in parent-related stress) with d=1.10. b) Child-related psychological stress CPP mothers experienced a significant decrease in child-related psychological stress pre- and post-intervention, which was also significantly different from mothers in the NC and CS groups (d=2.61 and 2.29 respectively). It is unclear whether the difference between the CPP and PPI groups was significant where d=1.45. NC and CS groups reported increases in child-related psychological stress that was not statistically significant. Further conditional latent difference score analysis was conducted to analyse whether treatment related changes in maternal psychological stress from baseline to post-intervention was associated with changes in basal cortisol activity from post-intervention to 1-year post intervention. The latent change score was regressed onto initial status to control for its effect and analysis was also controlled for pre-intervention levels of maternal parenting psychological distress variables. The only significant findings from this analysis related to CPP mothers, as it found in the CPP group a decrease in child-related psychological stress predicted a	**Overall validity score** + More precision on results in providing confidence intervals for differences between groups and a power calculation combined with no blinding.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
post-intervention (child m age 27.6 months) and 1 year post-intervention (child m 39.13 months). Outcome measures of interest are levels of maternal psychological parenting stress, measured using the Parenting Stress Index (PSI: Abidin 1997) completed at baseline and post-intervention and physiological stress, measured via cortisol levels in saliva samples, measured at baseline and 1 year post-intervention follow-up. Home-based and centre-based research sessions were conducted with mothers and their children, with demographic questionnaires completed by mothers during home-based interviews and self-report measures and saliva samples provided during lab-based sessions. The interventions CPP, and PPI took place in mothers own	• Type of abuse– Only mothers who were subject to CPS reports of neglect were analysed by this study. Neglect was defined by this study as caregiver who 'failed to provide for a child's basic needs for food, clothing, shelter, medical, care, adequate hygiene, or physical safety.' Incidents coded as neglect were 'inadequate supervision, maintaining unsanitary living conditions failing to seek medical care, or to provide adequate nourishment' (p9). In the maltreatment sample mothers were more likely to have experienced abuse (72%) than comparison mothers (28%) (x2 (1)=5.73<.001). Mothers in the maltreatment sample - 74% reported having experienced physical abuse; 66% reported sexual abuse; 55% reported experiencing neglect as a child. 30% reported experiencing 2 types of maltreatment, 54% reported experiencing 3 or more. • Looked after or adopted status– All infants living with biological mothers. • Unaccompanied asylum seeking, refugee or trafficked children– Not reported. **Sample size** Comparison numbers - Nonmaltreating comparison group	decrease in basal cortisol activity post intervention (B=.12, beta=.19; SE=.06, p<.05).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
homes. Structural equation models were constructed with AMOS software. Full-information maximum likelihood (FIML) was used to adjust for missing data. A latent difference score approach was used to analyse change in outcomes over time, and how the groups differed from each other, as well as the predictive link between psychological and physiological levels of stress. Time of day of saliva collection was controlled for, as cortisol levels in saliva decline throughout the day. **Country:** Not UK. US. **Source of funding:** Voluntary/charity - Spunk Fund Government - National Institute of Mental Health (MH54643).	n=52 Intervention numbers - CPP group n=44 PPI group n=34 Maltreating community standard comparison group n=27 Sample size - Maltreating group n=105 Nonmaltreating comparison group n=52. **Intervention** Describe intervention Two home-based preventative interventions, CPP and PPI, both designed to improve the mother-infant relationships and improve parenting in families that have been reported for maltreatment. **Delivered by** CPP and PPI are both delivered by masters levels therapists who participated in individual and group supervision on a weekly basis, with videotapes of sessions viewed by an independent third party to verify and maintain fidelity. **Delivered to** Mother-infant dyads. Duration, frequency, intensity, etc. Sessions delivered over 12 month period. **Key components and objectives of**		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	intervention CPP is based on the understanding that a caregiver's childhood experiences of parenting will affect their own. It uses a supportive and nondidactic approach to improving the mother-infant relationship by providing responsive developmental guidance and therapy that helps mothers to understand the effect of their pasts on their own parenting. It aims to improve their responsiveness, sensitivity and attunement to the needs of their child, and therefore begin to build a secure attachment between mothers and children. Beyond the mother-child relationship it also aims to help mothers build a supportive relational network. PPI in contrasts looks at current rather than past concerns. It involves didactic home-based parenting education and skills training, problem solving and relaxation techniques. It uses cognitive and behavioural techniques to address parenting skills deficits; limited personal resources, poor social supports and stresses in the home (those associated with maltreatment) while encouraging mothers to seek further education and employment and improving their social support networks. **Location/place of delivery** Mother's homes.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Describe comparison intervention In the maltreatment sample, there was a comparison with community standard services for maltreatment (no further detail given). There was also comparison with a sample who were not maltreating and therefore required no services. **Outcomes measured** Caregiver/parent health and wellbeing outcomes Maternal psychological and physiological parenting stress levels were measured using the Parenting Stress Index self-report questionnaire and assessments of saliva samples for cortisol levels. Results relating to cortisol levels have not been reported here. Parent-related stress (concerns about parental efficacy and competence; parental health and social isolation; relationship with others) was differentiated from child-related stress (parental perception of stress related to dealing with child's fluctuating mood, low adaptability and high demanding behaviour; difficulty in behaviour regulation and inability to reinforce parenting role).		

23. Winokur M, Ellis R, Drury I et al. (2014) Answering the big questions about differential response in Colorado: Safety and cost outcomes from a randomized controlled trial. Child Abuse and Neglect 39: 98–108

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: 1) To assess the impact on child safety outcomes of a family assessment response versus an investigation response assigned to children and families with a referral for child neglect or abuse 2) To examine the cost implications for child welfare agencies that implement a DR (Differential Response)-organized CPS (child protective services) system. **Methodology:** RCT including cluster. **Country:** Not UK. USA - Colorado. **Source of funding** Government. The Children's Bureau, Administration for Children and Families, US Department of Health and Human Services.	**Participants** Children and young people - children and families with a referral for child neglect or abuse Inclusion criteria for Family assessment response (FAR) cases: (1) mild to moderate general neglect, (2) educational neglect, (3) mild to moderate neglect from an injurious environment due to domestic violence, or (4) mild to moderate physical abuse. Exclusion criteria for FAR cases: Families with allegations of serious harm, sexual abuse, or suspicious child fatality were ineligible for FAR. Caregivers and families - children and families with a referral for child neglect or abuse Inclusion criteria for FAR cases: (1) mild to moderate general neglect, (2) educational neglect, (3) mild to moderate neglect from an injurious environment due to domestic violence, or (4) mild to moderate physical abuse. Exclusion criteria for FAR cases: Families with allegations of serious harm, sexual abuse, or suspicious child fatality were ineligible for FAR. **Sample characteristics** • Age - Not reported At baseline FAR families (Mean=5.9 years) had an	**Effect sizes** **Incidence of abuse and neglect** Safety outcomes by regression analysis (Table 3). 1. Referral within 365 days of initial referral: FAR=1407 (44%) vs. IR=820 (45%) (ns), total=2,227 (45%) 2. Assessment within 365 days of initial referral: FAR= 837 (26%) vs. IR=490 (27%) (ns), total=1,327 (27%) 3. High risk assessment (HRA) within 365 days of initial referral: FAR= 390 (12%) vs IR=243 (13%) (ns), total = 633 (13%) 4. Founded HRA within 365 days of initial referral: FAR=142 (4%) vs. IR=79 (4%) (ns), total=221 (4%) 5. Traditional Child welfare (CW) case opened after initial involvement: FAR=234 (7%) vs. IR=160 (9%) (ns), total=394 (8%) 6. Out-of-home (OOH) placement after initial involvement: FAR=188 (6%) vs IR=108 (6%) (ns), total=296 (6%). Survival analysis of the time-to-event data for predicting (Table 4) 1. The time to first referral: FAR and IR difference Hazard ratio (HR) 0.961, p=0.3816 (ns) 2. The time to assessment after the initial referral: FAR and IR difference Hazard ratio (HR) 0.975, p=0.6427 (ns) 3. The time to HRA after the initial referral: FAR and IR difference Hazard ratio (HR) 0.820, p=0.0100 (sig)	**Overall assessment of internal validity** + **Overall assessment of external validity** + **Overall validity score** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	'older' youngest child, on average, than did IR families (Mean =5.4 years). • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief– Not reported. • Disability– Not reported. • Long term health condition– Not reported. • Sexual orientation– Not reported. • Socioeconomic position– Not reported At baseline IR families (Mean=2.0) had more children in the home, on average, than did FAR families (M=1.8); IR families (Mean=1.7) had more caregivers, on average, than did FAR families (Mean=1.6). • Type of abuse– neglect and abuse. • Looked after or adopted status– All infants living with biological mothers. • Unaccompanied asylum seeking, refugee or trafficked children– Not applicable. **Sample size** Comparison numbers - Investigation response (IR), n=1,802 cases Intervention numbers - Family assessment response (FAR), n=3,194 cases	4. The time to founded HRA after the initial referral: FAR and IR difference Hazard ratio (HR) 0.932, p=0.5829 (ns) The predicted hazard rates for each of the significant predictor variables were the same for both FAR and IR tracks with no significant interactions between tracks and the predictors (raw data not presented here). FAR cases had a lower probability for referrals, assessments, high risk assessments, and founded high risk assessments over time than did IR cases. The FAR/IR difference was only statistically significant for predicting the probability of a high risk assessment (p = 0.01). FAR cases were 18% less likely to have an HRA, over time, than were IR cases.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Sample size - Overall sample size of 4,996 cases. **Intervention** **Describe intervention** In the differential response model, low- and moderate- risk families receive a family assessment response (FAR), a comprehensive assessment of family needs and strengths instead of a mal-treatment determination. Services are voluntary. **Delivered by** Not reported. **Delivered to** Families and children under referral for CAN. **Duration, frequency, intensity, etc.** Not reported. **Key components and objectives of intervention** Differential response also includes specified organisational processes in-cluding: enhanced screening, group supervision, family meetings and sup-port planning. The intervention also in-cludes rigorous assessment and use of evidence-based assessment tools, use of 'risk and goal statements' (p99)		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	and 'behaviourally-based safety and support plans'. **Content/session titles** Not reported. **Location/place of delivery** Not reported. **Describe comparison intervention** The control group received the investigation response (IR), a maltreatment determination with the possible provision of services (after opening a traditional child welfare case). **Outcomes measured** Incidence of abuse and neglect Safety outcomes from regression analysis as measured by: 1. Referral within 365 days of initial referral, 2. Assessment within 365 days of initial referral, 3. High risk assessment (HRA) within 365 days of initial referral, 4. Founded HRA within 365 days of initial referral, 5. Traditional Child welfare (CW) case opened after initial involvement, 6. Out-of-home (OOH) placement after initial involvement Safety outcomes from survival analysis as measured by the time-to-event data for predicting the time to the first referral, assessment, HRA, or founded HRA after the initial referral.		

Research question 16 – What is the impact of social and psychological interventions responding to child sexual abuse? (Prevention of recurrence, prevention of impairment)?

Research question 16 – Critical appraisal tables

1. Barbe RP, Bridge AJ, Birmaher B et al. (2004) Lifetime history of sexual abuse, clinical presentation, and outcome in a clinical trial for adolescent depression. Journal of Clinical Psychiatry 65: 77–83

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Aim of the study: The study aimed to determine what impact sexual abuse had on clinical depression and treatment outcome in depressed adolescents. **Description of theoretical approach?** No. The authors do not present a theory of change or logic model for any of the interventions and they do not hypothesise what impact history of sexual abuse will have on treatment outcome or why the intervention may work differently for different groups. **How was selection bias minimised?** Randomised. Method not reported. Participants were randomised to 1 of 3 treatment conditions but data from the systemic behavioural family therapy group were excluded from this analysis as only 1 participant in this group was found to have a history of sexual abuse. The authors note	**Was the exposure to the intervention and comparison as intended?** Yes. Some sessions were videotaped and rated by experts which the authors report showed that all three treatments were '… delivered with fidelity and were distinct from each other' (p78). **Was contamination acceptably low?** Not reported. Information on contamination is not provided. **Did either group receive additional interventions or have services provided in a different manner?** Partly. Seven participants were removed from the protocol and referred to 'open treatment' because they had failed to make symptomatic progress. It is unclear what 'open treatment' refers to. There is no other indication that any of the treatment groups received any other additional interventions or received services in a different manner.	**Does the study's research question match the review question?** Partly. The study aimed to determine what impact sexual abuse had on clinical depression and treatment outcome in depressed adolescents. Sexual abuse is therefore treated as a moderating variable. However, the reviewing team decided that this study was still relevant to the review question as it gathers data about the effectiveness of cognitive behavioural therapy for abused compared to non-abused individual. Only data relating to the interaction between treatment effects and abuse history have been extracted. **Has the study dealt appropriately with any ethical concerns?** Yes. The study protocol was approved by an institutional review board and informed consent was sought from participants and their families.	**Overall assessment of internal validity:** A key limitation is the referral of participants not showing progress to 'open treatment'. Also, the authors note that randomisation included 'balancing' for number of parents in household, gender, suicidality. It is unclear whether this refers to a priori stratification, or a post hoc adjustment to randomisation, which would in turn negate the randomisation process. **Overall assessment of external validity:** ++ **Overall validity rating:** -

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
that randomisation included 'balancing' for number of parents in household, gender, suicidality. It is unclear whether this refers to a priori stratification, or a post hoc adjustment to randomisation, which would in turn negate the randomisation process. **Was the allocation method followed?** Not reported. Method of allocation and concealment are not reported. **Is blinding an issue in this study?** Part-blinding. Due to the nature of the intervention it would not have been possible to blind participants or providers; however the authors report that assessments were conducted by clinical interviewers who were blinded to group assignment. **Did participants reflect target group?** Yes. An acceptable number of eligible adolescents agreed to participate (87.7%). The authors report that participants were moderately depressed (mean Beck Depression Inventory score of 24.1 (SD=8.1) and they note that participants were '… nonpsychotic, nonbipolar, and without ob-	**Were outcomes relevant?** Yes. The study focused on the impact which history of sexual abuse had on clinical depression at baseline and treatment outcome for depression. Presence of depression was used as the primary outcome. **Were outcome measures reliable?** Yes. The authors note that all measures have established reliability and validity; however they do not present data in relation to this. The majority of measures were based on child and parental self-report with the exception of major depression which was also assessed by an interviewer blinded to group assignment, and functional impairment which was only assessed by an interviewer blinded to group assignment. History of sexual abuse was determined by a yes/no question which the authors themselves note is a limitation, meaning that the impact of frequency or severity of abuse could not be investigated. **Were all outcome measurements complete?** Yes. All planned data was gathered; however only the primary outcome of	**Were service users involved in the study?** No. Service users involved as participants only. No indication that service users were involved at the design stage or in the interpretation of findings. **Is there a clear focus on the guideline topic?** Yes. The study focuses on the effect which sexual abuse has on clinical depression and treatment outcome in depressed adolescents. **Is the study population the same as at least 1 of the groups covered by the guideline?** Partly. Participants were adolescents between the ages of 13 and 18. Only 11 adolescents were identified as having experienced sexual abuse out of a total sample of 107 however the study has been included in the NCCSC review because findings are reported on the basis of history of sexual abuse status. NB. Although participants were randomised to 1 of 3 treatment conditions this study only reports on those randomised to either cognitive-behavioural therapy or nondirective supportive therapy (n=72) due to the fact that only one participant assigned to systemic behavioural	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
sessive-compulsive disorder, eating disorder, substance abuse, or ongoing physical or sexual abuse' (p78). The Guideline Committee may therefore wish to bear in mind that the findings of this study are not generalisable to younger children or adolescents with severe psychological symptoms. **Were all participants accounted for at study conclusion?** Not reported. It is not clear whether all participants took part in all follow-up assessments (7 in total, in addition to baseline) as the authors only report the percentage which took part in the 6 week (sixth session) and 24 months post-treatment assessments. Of the total sample, 87.8% (n=94) completed an interview and the Beck Depression Inventory at the 6 week/session time point. 92.5% (n=99) completed the 24 months post-treatment interview, and 90.7% (n=97) completed the Beck Depression Inventory at this time point. Drop-out rates were acceptable and did not differ significantly by treatment group. The authors also report that there was no significant difference in drop-out rates between participants with a history of sexual abuse and those	presence of DSM-III-R major depression was measured at both time points. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. All groups were followed-up for the same length of time. Final assessments took place 24 months after completion of treatment. **Was follow-up time meaningful?** Yes. Assessments took place at baseline, at the sixth week/session, at treatment completion, every 3 months during the first year post-treatment, and at the 24 months post-treatment point. These would be sufficient to capture the more immediate effects of treatment as well as the intermediate effects. **Were the analytical methods appropriate?** Partly. Analytical measures were appropriate. However, only one data was analysed in terms of the interaction between treatment group and abuse history – the variables of interest to this review.	therapy was determined to have a history of sexual abuse. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Not reported. The context in which treatments and assessments took place is not reported by the authors. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Response (prevention of impairment). The study aims to measure the impact of sexual abuse history on clinical depression and treatment outcome in a randomised controlled trial evaluating three types of therapy. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Outcomes included rate of major depression, rate of decline in depressive symptoms and psychiatric hospitalisation. **Does the study have a UK perspective?** No. The study was conducted in the USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
without. The authors also report that seven participants were referred to 'open treatment' and excluded from the research protocol due to the fact that they '… continued to meet criteria for major depression, had a Beck Depression Inventory score persistently higher than or equal to 13, and had failed to make symptomatic progress' (p78). A slightly higher proportion of these participants came from the comparison interventions than from the cognitive behavioural therapy intervention. The authors report that rates of referral did not differ by treatment group or by history of sexual abuse status. It appears that adolescents referred to 'open treatment' were still included in the final analysis.	**Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Not reported. Pre-treatment analysis focused on comparing sexually abused children to those who had not experienced sexual abuse, rather than assessing any differences between exposure and comparison. The authors report that these 2 groups were similar on the majority of variables however rates of maternal current major depression were significantly higher in the sexually abused group (78% vs. 39%, Fisher exact test p=.04). **Was intention to treat (ITT) analysis conducted?** Unclear. Data appear to be based on all 72 randomised participants, but the authors do not describe undertaking an intention to treat analysis. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. The authors do not provide a power calculation. The sample size is acceptable. **Were the estimates of effect size given or calculable?** Partly.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Odds ratios and phi values are presented in some instances but the reporting of these is not consistent. **Was the precision of intervention effects given or calculable?** Yes. P values and confidence intervals are presented as appropriate. **Do conclusions match findings?** Yes.		

2. Carpenter J, Jessiman T, Patsos D et al. (2016) Letting the future in: A therapeutic intervention for children affected by sexual abuse and their carers. London: NSPCC

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim Study aim: Research questions: '1. What are the outcomes for children and young people affected by sexual abuse of providing Letting the Future In in NSPCC services? 2. What is the cost-effectiveness of this service? 3. What is the effectiveness of the support intervention received by the 'safe' carers?' (p10). **Description of theoretical approach?** Yes.	**Was the exposure to the intervention and comparison as intended?** Partly. Important to note that some intervention sessions happened after T2 data collection point (key data collection point). **Was contamination acceptably low?** Yes. **Did either group receive additional interventions or have services provided in a different**	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Yes. Research approved by Research Ethics Committee of the NSPCC and the Research Ethics Committees of the Universities of Bristol and Durham. Informed consent sought from children and non-abusing parents.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Clear description of theoretical origins of the model, which is rooted in Bannister's (2003) Recovery and Regeneration model, which is influenced by psychodrama, play therapy and attachment theory. Model is described as 'deliberately multi-theoretical' (p21), and uses diverse constructs to build a guide for practitioners (not a manualised treatment) which has: - a value base - an underpinning knowledge base - a skills base. **How was selection bias minimised?** Randomised. Randomised by clinical trials centre in the Netherlands. **Was the allocation method followed?** Yes. **Is blinding an issue in this study?** No blinding. Wait list control - no blinding. **Did participants reflect target group?** Yes. **Were all participants accounted for at study conclusion?** Yes. Data completed at T2 for 72% of intervention group and 73% of	**manner?** No. **Were outcomes relevant?** Yes. **Were outcome measures reliable?** Partly. All outcome measures were via self-report - no clinician assessment or observation. **Were all outcome measurements complete?** Partly. Some participants left intervention or were lost to follow up. Missing data were imputed. **Were all important outcomes assessed?** Partly. Lack of clinical assessment at follow-up noted as a limitation by study authors. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Partly. Majority of analyses use T2 data - gathered at 6 months. This is a relatively short follow-up time.	**Were service users involved in the study?** No. Service users involved as participants only. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children who have been sexually abused and their non-abusing carer. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to Response. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
control group. Missing data imputed.	**Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. No significant differences between intervention and control groups at baseline. **Was intention to treat (ITT) analysis conducted?** Yes. Missing data were imputed. Study states that: 'Multiple imputation in this case was the statistical prediction of the missing score based on T1 score, demographics and variables suggested by previous literature to affect treatment outcomes, the nature of abuse (penetration or attempted penetration) and intra- or extra-familial abuse. Five imputations were run and a pooled estimate used' (p67). **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Partly. Study's power calculation estimates that 210 participants required - study achieved sample of 242. However, authors note that not all the children in the sample met the clinical threshold on the Trauma Symptoms Checklist for Children at baseline, and so were		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	unable to show a positive result on the primary outcome (reduction in 'caseness' as measured by TSCC). This effectively meant that the study was underpowered. **Were the estimates of effect size given or calculable?** No. Authors report test statistics and p values only. **Were the analytical methods appropriate?** Partly. Reporting of results of McNemar's test for primary outcomes is unclear. **Was the precision of intervention effects given or calculable? Were they meaningful?** No. **Do conclusions match findings?** Yes.		

3. Danielson CK, McCart MR, Walsh K, et al. (2012) Reducing substance use risk and mental health problems among sexually assaulted adolescents: A pilot randomized controlled trial. Journal of Family Psychology 26: 628–35

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim To evaluate the differential efficacy of Reduction through Family Therapy (RRFT) and treatment as usual (TAU) in reducing substance use (SU) problems (including early initiation), mental health symptoms, and risky sexual behavior among adolescent CSA victims. **Description of theoretical approach?** Yes. Several theories are incorporated into RRFT via multiple intervention strategies. First, RRFT is guided by ecological theory, which proposes that an adolescent's behavior is influenced by multiple social and environmental contexts, including the family, peer network, school and community (Bronfenbrenner 1979). RRFT, like other ecological models (e.g., MST), adopts a family-based approach to intervention and encourages therapists to intervene in multiple social systems. As part of this ecological model, Strategic Family Therapy (e.g., Haley 1976) is utilized to help the family define problems (in behaviourally specific terms – and from the perspective of both of the adolescent and the caregiver) and work together to	**Was the exposure to the intervention and comparison as intended?** Not reported. **Was contamination acceptably low?** Not reported. **Did either group receive additional interventions or have services provided in a different manner?** Not reported. **Were outcomes relevant?** Partly Yes to 1 aspect of the review that is seeking to understand what evidence there is about the health and wellbeing of young people who have experienced child abuse. **Were outcome measures reliable?** Partly Most of the outcomes are based on robust pre-tested scales and measures. However, it should be noted there were some limitations with the study: baseline differences existed between RRFT and TAU across most variables. This was likely due to the small sample size and the unrestrictive inclusion	**Does the study's research question match the review question?** Yes. The study is evaluating the impact of Risk Reduction through Family Therapy (RRFT) for reducing substance use risk and trauma-related mental health problems among sexually assaulted adolescents. **Has the study dealt appropriately with any ethical concerns?** Yes. Eligible adolescents and their (non-offending) caregivers were approached by a researcher to solicit participation. Of the eligible families, 91% agreed to participate (see Figure 1). Legal guardians provided consent and youth provided assent. Following consent, participants were randomly assigned to condition using a computerised blocked randomisation method. Participants were compensated for completing each assessment but were not paid to attend treatment. All procedures were approved by MUSC's Institutional Review Board, and a certificate of confidentiality was obtained. Adequate consideration	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** + **Overall validity rating:** + + (moderate) external validity (related to narrow match of study outcomes relative to review scope) but ++ internal validity = + moderate overall validity.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
solve those problems. The family is involved across all 7 treatment components. Mowrer's Two-Factor Theory (Mowrer 1960) and negative reinforcement theory (Baker et al. 2004) also are applied in targeting PTSD, SU, and their overlap in RRFT. Additional information about the RRFT model can be found in Danielson et al. (2010a). **How was selection bias minimised?** Randomised. Following consent, participants were randomly assigned to condition using a computerised blocked randomisation method. **Was the allocation method followed?** Partly. Adherence to RRFT was assessed via review of randomly selected audiotaped sessions (two per client per month) by the first author. Given that the control condition was 'treatment as usual' (i.e., not one specific protocol), adherence was not monitored systematically in this condition. RRFT treatment adherence also was monitored through weekly individual supervision with the treatment developer. Further, participants	criteria, which permitted both substance using and non-substance using youth to participate (i.e., to capture a 'real world' adolescent CSA sample). Although this heterogeneity in symptom presentation among study participants is representative of the multiple trajectories of adolescents with a CSA history (Danielson et al. 2010b), which drove study design decisions to focus on this high risk population rather than a particular diagnosis, future studies will need to ensure equality on key variables across the two conditions. Similarly, dosage differences existed between the 2 conditions, where RRFT youth received more sessions on average than TAU youth. Analyses indicated that dosage was not significantly associated with any of the outcome variables. Nonetheless, an important future direction will involve use of an attention-matched control condition, as well as a much larger sample size to allow for an examination of mediators and/or moderators that may speak to RRFT's 'active ingredients.' Finally, while the flexibility of offering treatment outside of the office can be viewed as a strength and rep-	was also given to the administration of the treatment and quality assurance. Participants were treated by clinical psychology graduate students completing a predoctoral internship. Therapy was delivered through both an outpatient clinic and an outreach programme offered at the same clinic (for families without transportation; 4 families in RRFT condition and 3 in TAU condition). The first author, a licensed clinical psychologist and RRFT developer, supervised all RRFT cases. TAU therapists were supervised by other licensed psychologists in the clinic. RRFT therapists received didactic training on the intervention prior to implementation. Adherence to RRFT was assessed via review of randomly selected audiotaped sessions (2 per client per month) by the first author. Given that the control condition was 'treatment as usual' (i.e., not one specific protocol), adherence was not monitored systematically in this condition. RRFT treatment adherence also was monitored through weekly individual supervision with the treatment developer. Further, participants were asked to complete an RRFT adherence checklist immediately following	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
were asked to complete an RRFT adherence checklist immediately following each session, which was then reviewed. **Is blinding an issue in this study?** Blinding. Following consent, participants were randomly assigned to condition using a computerised blocked randomisation method. Participants then completed a baseline assessment, which included the measures noted below and a urine drug screen. Assessment measures were re-administered at post-treatment, 3-, and 6-month follow-up using a match-timing design, where completion of treatment by a RRFT youth triggered the post-treatment assessment (and timing for the 3-month and 6-month posttreatment assessments) for that youth and his/her matched pair in the TAU condition. **Did participants reflect target group?** Partly. Total 102 were assessed foe eligibility: 72 were excluded and 30 were included. Of the excluded, 69 did not meet the inclusion criteria, and 3 refused to participate.	resentative of the real world barriers faced by many clients, this introduces another layer of heterogeneity within the sample and thus 'noise' when drawing conclusions about treatment efficacy. **Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Yes. Assessment measures were re-administered at post-treatment, 3-, and 6-month follow-up using a match-timing design, where completion of treatment by a RRFT youth triggered the post-treatment assessment (and timing for the 3-month and 6-month posttreatment assessments) for that youth and his/her matched pair in the TAU condition. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes.	each session, which was then reviewed. **Were service users involved in the study?** Yes. Participants were attendees of a weekly mental health service. Participants were recruited through an urban clinic specialising in the treatment of trauma. **Is there a clear focus on the guideline topic?** Partly. This study is a specific intervention supporting young people who have experienced child abuse and neglect. However, it does not set out to help explain what may prevent future abuse. Rather its focus is on supporting young people who have experienced child abuse recover from their trauma. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. The study covers young people (under 18) who have experienced abuse or neglect and/or their caregivers and families. **Is the study setting the same as at least 1 of the settings covered by the guideline?**	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Were all participants accounted for at study conclusion? Yes. 30 were included in the study - 15 to the treatment group and 15 to the control group. All results are reported of the 30 indicating no attrition or dropout.	Table 1 includes descriptive data for each outcome. The MRMs yielded significant condition effects on the UCLA PTSD-P, CDI, BASC-Internalising, TLFB, FES Cohesion (adolescent and parent-report), and FES Conflict (adolescent-report) scales (see Table 2), reflecting greater impairment among RRFT youth relative to TAU youth at baseline. These baseline differences need to be considered when interpreting the results pertaining to between-group differences on change over time. **Was intention to treat (ITT) analysis conducted?** Yes. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. **Were the estimates of effect size given or calculable?** Yes. Standardised beta coefficients were reported in Table 2. **Were the analytical methods appropriate?** Yes. Data were comprised of 4 repeated measurements (level-1)	Yes. The setting is a secondary health centre. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes Area e: Specific interventions as part of a package of support for children and young people who have experienced child abuse and neglect. The treatment Risk Reduction through Family Therapy (RRFT) has been aimed at reducing substance use risk and trauma-related mental health problems among sexually assaulted adolescents. This includes family therapy. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Partly. One but not all. The study focuses on young people's health and wellbeing. **Does the study have a UK perspective?** No. The study is set in the USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	nested within 30 families (level-2), yielding a two-level Mixed-Effects Regression Model (MRM). The TLFB score represented the number of days with self-reported SU over the previous 90 days and was modeled as a count-distributed (i.e., Poisson) outcome. A Sexual Partners variable was modelled as a dichotomous (i.e., Bernoulli) outcome indicating whether the adolescent had any new consensual sexual partners over the previous 90 days. The remaining outcomes were modeled as continuous variables. Slopes were modeled using linear polynomials computed from assessment dates (Singer & Willett 2003). Treatment condition was coded such that RRFT=0 and TAU=1. MRMs were performed using HLM software (v. 6.08; Raudenbush et al. 2004), with restricted maximum likelihood estimation for the continuous outcomes and Laplace estimation for the Poisson and Bernoulli outcomes. Specification of random effects was based on the likelihood ratio test for the continuous outcomes and the Wald test for variance components for the Poisson and Bernoulli outcomes (Singer & Willett 2003). The Wald		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	test statistic for significance testing was computed using asymptotic standard errors. According to this model specification, the intercept and slope terms represent the average baseline score and monthly rate of change for youth in the RRFT condition. The condition and condition × linear terms represent the difference between TAU and RRFT at baseline and in the monthly rate of change. Variables representing number of sessions and treatment length were initially included in the MRMs to control for treatment intensity. Conclusions did not differ when intensity was controlled; therefore, results are presented for models without the covariates. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. Table 2 reports p values (with CI) for each beta coefficient. **Do conclusions match findings?** Yes. The conclusions are consistent with stated hypotheses, adolescents who received RRFT reported reduced SU and improve-		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	ments in SU risk factors (e.g., increased family cohesion). Also as expected, participants in both conditions experienced reductions in PTSD and depression symptoms, although greater reductions were found for adolescents in the RRFT condition with regard to parent-reported PTSD, as well as adolescent-reported depression and internalising symptoms. However, randomization failed to prevent inequality at baseline across the 2 conditions.		

4. Foa EB, McLean CP and Capaldi S (2013) Prolonged exposure vs supportive counseling for sexual abuse-related PTSD in adolescent girls: A randomized clinical trial. Journal of the American Medical Association 310: 2650–7

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Aim of the study: The study aims to evaluate the impact of counsellor-delivered prolonged exposure therapy in comparison to supportive counselling for adolescent girls with sexual abuse related post-traumatic stress disorder. **Description of theoretical approach?** No. The authors do not outline the theories which underpin prolonged exposure therapy. They note that although the effects of this therapy have been extensively studied in adults it is not	**Was the exposure to the intervention and comparison as intended?** Not reported. The authors do not report any changes to either the intervention or comparison during the course of the study. Treatment completion was defined as receipt of at least 8 of the 14 treatment sessions for both intervention and comparison conditions. Three participants assigned to the intervention (9.7% and 5 assigned to the comparison condition (16.6%) failed to complete	**Does the study's research question match the review question?** Yes. The study aims to evaluate the impact of counsellor-delivered prolonged exposure therapy in comparison to supportive counselling for adolescent girls with sexual abuse related post-traumatic stress disorder. **Has the study dealt appropriately with any ethical concerns?** Yes. The study was approved by the University of Pennsylvania in-	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** ++ **Overall validity rating:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
usually offered to adolescents due to concerns that it can only be safely provided to participants who have 'mastered' the necessary skills to cope with exposure and can exacerbate post-traumatic stress disorder symptoms. **How was selection bias minimised?** Randomised. Permuted block procedure. **Was the allocation method followed?** Not reported. Method of allocation and concealment are not reported. **Is blinding an issue in this study?** Part blinding. Due to the nature of the intervention it would not have been possible to blind participants or providers, however the investigators who conducted assessments were blinded to group assignment. **Did participants reflect target group?** Partly. An acceptable number of eligible individuals agreed to participate, however a number of eligible participants were not randomised because they refused any treatment at all, wanted to take part in group treatment only or because there were	treatment. The authors report adherence to the essential treatment components as 90.8% for the intervention and 90.5% for the comparison condition. **Was contamination acceptably low?** Not reported. The authors do not provide detail on contamination. **Did either group receive additional interventions or have services provided in a different manner?** Not reported. There is no indication that either group received additional services. **Were outcomes relevant?** Yes. The primary outcome was post-traumatic stress disorder symptom severity. Secondary outcomes included post-traumatic stress disorder diagnosis, depression severity, and functioning and these were measured directly. **Were outcome measures reliable?** Yes. All outcome measures relating to treatment efficacy had established reliability and validity although data to support this is not always presented. The authors report that the primary outcome measure (Child PTSD Symptom	stitutional review board and informed consent/assent was given by participants and their guardians. **Were service users involved in the study?** No. Service users involved as participants only, no indication of involvement at design stage or in interpretation of findings. **Is there a clear focus on the guideline topic?** Yes. The study evaluates the effects of an intervention on post-traumatic stress disorder in sexually abused adolescent girls. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Sexually abused adolescent girls. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. The study was conducted in a rape crisis centre. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Response (prevention of impairment). The study evaluates the effects of an	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
problems obtaining parental consent because of custody issues. Participants with less than three months since trauma were also excluded. Formal inclusion and exclusion criteria were '…current suicidal ideation with intent, uncontrolled bipolar disorder, schizophrenia, conduct disorder, pervasive developmental disorder, initiation of psychotropic medication within the previous 12 weeks, and current inpatient psychiatric treatment. Adolescents with substance use or suicidality without imminent threat were not excluded' (p2651). All participants were female, over half of whom were Black and 6 participants with late stage pregnancies were also excluded. **Were all participants accounted for at study conclusion?** Yes. Rates of intervention completion and availability of outcome data were acceptable and comparable between groups.	Scale–Interview) has '… excellent internal consistency (Cronbach α = .83-.89) and test-retest reliability (Cronbach α=.84-.86)…' as well as high convergent validity and discriminant validity (p2651). A mixture of self-report and clinically judged ratings are used. Participant's expectations regarding treatment were also assessed at the first session, and was measured using a scale which does not appear to be published, however this data was not used in analysis. **Were all outcome measurements complete?** Yes. All data was collected as planned; however there were some missing data which the researchers dealt with by using multiple imputation. Pattern mixture modelling showed that missing data did not affect any of the results. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. Both groups were followed-up for an equal length of time.	intervention on sexual abuse related post-traumatic stress disorder in adolescent girls. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. The primary outcome was post-traumatic stress disorder symptom severity. Secondary outcomes included post-traumatic stress disorder diagnosis, depression severity, and functioning. **Does the study have a UK perspective?** No. The study was conducted in the USA.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Was follow-up time meaningful? Yes. Follow up took place mid-treatment, post-treatment and at three, six and 12 months post-treatment. However, 12 months may not have been sufficient to detect longer-term benefits or harms. NB. Data are not reported for mid-treatment or 3 and 6 month's post-treatment assessments. **Were the analytical methods appropriate?** Yes. Continuous data were analysed using piecewise linear mixed models. Dichotomous data were analysed using generalised linear mixed models. The authors report that to account for the possibility of type 1 errors the Benjamini-Hochberg-Yekutieli procedure was used to correct p values. NB This article only reports the corrected values. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. There were no significant differences between groups on any demographic variables or on any of the outcome measures (all p>.17). **Was intention to treat (ITT) analysis conducted?** Yes.		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Was the study sufficiently powered to detect an intervention effect (if one exists)? Yes. The authors report a power calculation showing that the study had greater than 0.95 power to detect medium effect sizes in differences between group means. **Were the estimates of effect size given or calculable?** Partly. Cohen's d are reported for the primary outcome of post-traumatic stress disorder. Effect sizes for secondary outcomes are not provided. **Was the precision of intervention effects given or calculable?** Yes. Confidence intervals and p values are provided. **Do conclusions match findings?** Yes.		

5. Goldman Fraser J, Lloyd SW, Murphy RA et al. (2013) Child exposure to trauma: Comparative effectiveness of interventions addressing maltreatment. Comparative Effectiveness Review 89: 1–161

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. The review seeks to evaluate the comparative effi-	**Inclusion of relevant individual studies?** Yes. Includes interventions for sexually abused children.	**Does the study's research question match the review**	**Overall assessment of internal validity:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
cacy and effectiveness of psychosocial and pharmacological interventions that address child well-being and/or promote positive child welfare outcomes (safety, placement stability, and permanency) for maltreated children ages birth to 14 years. **Adequate description of methodology?** Yes. **Rigorous literature search?** Yes. Four bibliographic databases searched, search of grey and unpublished literature, and trial registries.	**Study quality assessed and reported?** Yes. Study quality assessed for risk of bias and graded as 'High', 'Medium' or 'Low' risk of bias. **Do conclusions match findings?** Yes.	**question?** Yes. Study is applicable to both review questions 15 and 16. **Has the study dealt appropriately with any ethical concerns?** Not reported. **Were service users involved in the study?** Not reported. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children aged 0 to 14 years exposed to maltreatment. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Response following maltreatment. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. No data on abuse recurrence.	**Overall assessment of external validity:** + **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Does the study have a UK perspective? No. Studies conducted in the USA, the UK and Canada	

6. Leenarts, LE, Diehle J, Doreleijers TA et al. (2013) Evidence-based treatments for children with trauma-related psychopathology as a result of childhood maltreatment: A systematic review. European Child and Adolescent Psychiatry 22: 269–83

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Partly. The review question is clear and is relevant to the NCCSC guideline; however, the review lacks detail overall, i.e. in relation to direction of effect, significance values and there is a lack of clarity in relation to whether the included studies evaluate treatments which are psychotherapeutic or cognitive behavioural. **Adequate description of methodology?** No. On the whole, the review includes little detail on methodology, e.g. how the process by which the authors came to the three 'clusters' of treatment type. The findings are mostly narrative and no meta-analysis has been conducted. The authors only report on post-traumatic stress disorder related outcomes which could result in bias and it is not clear whether the search strategy included outcomes. Significance	**Inclusion of relevant individual studies?** Yes. The included studies are appropriate to address the review question as set out by the authors and are clearly relevant to the NCCSC work. **Study quality assessed and reported?** Yes. The authors assessed risk of bias in relation to seven domains using a Cochrane Collaboration tool (Higgins et al. 2011) and disagreements were solved by consensus. Although these ratings are presented in table form the individual ratings do not appear to have been combined in any way and it is therefore difficult to clearly understand the limitations of each study. Similarly, the narrative findings in relation to each type of treatment do not consistently include details on the quality of the studies.	**Does the study's research question match the review question?** Yes. The objective of the review is to '… systematically evaluate psychotherapeutic treatments for children exposed to childhood maltreatment and to describe treatments which focus on the above mentioned broad range of psychopathological outcomes' (p270). The review included studies focusing on treatment for sexually abused children and has been included on the basis that effect sizes for these are reported separately. **Has the study dealt appropriately with any ethical concerns?** Not reported. The review authors do not record whether the research protocols of individual studies were approved by institutional review boards or whether participants gave informed consent.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall assessment of validity:** -

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
levels are not provided and although it appears that the included effect sizes are Cohen's d the reviews authors do not make this clear or report whether these were calculated by themselves or reported by the original studies. In addition, where between group differences were detected, the direction of effect is not reported. **Rigorous literature search?** Partly. The authors appear to have searched an appropriate number of relevant databases, however there is no indication that citation searching was conducted. In addition, full search strategies are not reported which means that it is difficult to determine whether this was comprehensive (although the authors report that both controlled vocabulary and free text terms were used.	**Do conclusions match findings?** Yes.	**Were service users involved in the study?** No. Service users appear to be involved as participants of individual studies only. There is no indication of involvement at the design stage or in the interpretation of findings. **Is there a clear focus on the guideline topic?** Yes. The study focuses on psychotherapeutic treatments provided to children who have experienced maltreatment in order to determine their impact on psychopathological outcomes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children and young people between the ages of 6 and 18 years who had experienced maltreatment. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Not reported. The review authors do not report whether the individual studies recorded settings or contexts. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Response	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		(prevention of impairment). The review focuses on psychotherapeutic treatments which are designed to address the psychopathological effects of maltreatment. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. The review focuses on the 'broad range of psychopathological outcomes' such as anxiety, post-traumatic stress disorder, suicidal ideation, and substance abuse, which can result from maltreatment. **Does the study have a UK perspective?** No. Only one of the included studies originates from the UK, and the majority were conducted in the USA. The review was conducted by Dutch researchers.	

7. Macdonald G, Higgins J, Ramchandani P et al. (2012) Cognitive-behavioural interventions for children who have been sexually abused: A systematic review. Campbell Systematic Reviews 14: 111

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. The review question is clearly focused and is relevant to the NCCSC review work. However, the detail provided on	**Inclusion of relevant individual studies?** Yes. The included studies are appropriate to address the review question as set out by the authors and are clearly relevant to the NCCSC work.	**Does the study's research question match the review question?** Yes. The objective of the review is to '… assess the efficacy of cognitive-behavioural approaches (CBT) in addressing the	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
the reviews inclusion and exclusion criteria are quite minimal, e.g. search dates. **Adequate description of methodology?** Yes. The approaches used to analyse the data are explained fully and are justifiable. Hedge's g effect sizes are used. Two studies were randomised by group rather than individual. The authors estimate the possible inflation of effect size due to clustering effects, however it is not clear whether the study's effect size was adjusted to take this in to account. **Rigorous literature search?** Yes. The authors searched an appropriate number of relevant databases and some citation searching appears to have been carried out. The search strategies are comprehensive and relevant to the focus of the review and both controlled vocabulary and free text terms are used.	**Study quality assessed and reported?** Yes. The authors assessed risk of bias in relation to six domains using a Cochrane Collaboration tool (Higgins 2008) and these are clearly presented in table form. However these individual ratings do not appear to have been combined in any way and it is therefore not clear what the overall risk of bias for each study is. The authors note that the included studies were generally of a low quality and that this is exacerbated by poor reporting; but the way in which the authors arrive at this conclusion is not transparent. **Do conclusions match findings?** Yes.	immediate and longer-term sequelae of sexual abuse on children and young people up to 18 years of age' (p6). **Has the study dealt appropriately with any ethical concerns?** Not reported. The review authors do not record whether the research protocols of individual studies were approved by institutional review boards or whether participants gave informed consent. **Were service users involved in the study?** No. Service users involved as participants only, no indication of involvement at design stage or interpretation of findings. **Is there a clear focus on the guideline topic?** Yes. The study focuses on the use of cognitive behavioural therapy in the treatment of sexual abuse related sequelae in children and young people up to the age of 18. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children and young people up to the age of 18 with experience of sexual abuse.	**Overall assessment of validity:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Is the study setting the same as at least 1 of the settings covered by the guideline? Not reported. The review authors do not report whether the individual studies recorded settings or contexts. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Response (prevention of impairment). Cognitive behavioural therapy to address the sequelae of sexual abuse. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Primary outcomes specified by the review authors are described as children's psychological functioning (e.g. anxiety, depression and post-traumatic stress disorder); and children's behavioural problems (e.g. externalising and sexualised behaviour). Secondary outcomes are summarised as 'future offending behaviours' such as sexual offending, delinquency and criminal offending; and outcomes relating to the quality of parenting and parent-child relationships which the authors categorise as	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		'parental skills and knowledge' (included behavioural management skills, parental emotional reactions, knowledge of child sexual abuse and possible consequences of this); 'belief in their child's story'; and understanding of child behaviour and psychological problems. NB. The review does not report on 'future offending behaviours' and it is therefore not clear if any of the included studies report on this type of outcome. **Does the study have a UK perspective?** No. The trials which the authors report on were conducted in the USA (n=9) and Australia (n=1). The review was conducted by researchers in Northern Ireland.	

8. Parker B and Turner W (2013) Psychoanalytic/psychodynamic psychotherapy for children and adolescents who have been sexually abused: A systematic review. Campbell Collaboration 9(13)

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Appropriate and clearly focused question? Yes. The review question is clearly focused and is relevant to the NCCSC review work. However, the detail provided on the reviews inclusion and exclusion criteria are quite minimal e.g. search dates.	**Inclusion of relevant individual studies?** N/A The review did not identify any studies which were eligible for inclusion, however the inclusion and exclusion criteria used by the reviewers were appropriate both for their own review question and that of the NCCSC with the exception that studies which used	**Does the study's research question match the review question?** Yes. The objective of the review is to '… assess the effectiveness of psychoanalytic/psychodynamic psychotherapeutic approaches in treating the effects of sexual abuse (psychologically	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** ++ **Overall assessment of validity:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Adequate description of methodology? Yes. The authors provide a good level of detail in relation to search strategies, inclusion and exclusion criteria and data collection and analysis methods. **Rigorous literature search?** Yes. An appropriate number of relevant databases were searched and citation searching was also conducted. The search strategies are comprehensive and relevant to the focus of the review and both controlled vocabulary and free text terms were used.	an active comparison group were excluded. **Study quality assessed and reported?** Unclear. The review did not identify any studies which were eligible for inclusion. **Do conclusions match findings?** Yes.	and in terms of behaviour and social functioning) in children and adolescents' (p13). **Has the study dealt appropriately with any ethical concerns?** Not reported. The review protocol does not appear to include consideration of these issues. **Were service users involved in the study?** Not reported. The review protocol does not appear to include consideration of this issue. **Is there a clear focus on the guideline topic?** Yes. The study focuses on the use of psychoanalytic/psychodynamic approaches to the treatment of sexual abuse related sequelae in children and adolescents. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children and young people up to the age of 18 with experience of sexual abuse. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Not reported. The review protocol does	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		not include criteria relating to settings or context. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Response (prevention of impairment). Psychoanalytic, psychodynamic or psychotherapeutic treatment approaches to address the sequelae of sexual abuse. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. Primary outcomes specified in the review protocol are; post-traumatic stress disorder, depression, sexualised behaviour, aggression or conduct problems, and self-harm. The protocol specifically includes suicide as an adverse outcome. Secondary outcomes are conceptualised as; symptoms and/or psychiatric diagnosis, 'measures of underlying processes (relevant to psychoanalytic/psychodynamic psychotherapy)', psychosocial functioning, service use, and 'other' (p16). **Does the study have a UK perspective?** Unclear. The review did not identify any studies which	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		were eligible for inclusion. The review was conducted by authors based in England.	

9. Shirk SR, DePrince AP, Crisostomo PS et al. (2014). Cognitive behavioral therapy for depressed adolescents exposed to interpersonal trauma: An initial effectiveness trial. Psychotherapy 51: 167–79

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim To assess the feasibility, acceptability and initial impact of a modified CBT protocol (based on 1 previously found to be effective, the Adolescent Mood Project protocol) for adolescent depression (m-CBT) designed to treat adolescents with history of interpersonal trauma. **Description of theoretical approach?** Yes. Modification of CBT in light of evidence that effectiveness in treating depression can be diminished when there is co-existing history of sexual abuse. **How was selection bias minimised?** Randomised. Randomisation stratified by gender.	**Was the exposure to the intervention and comparison as intended?** Yes. Study reports that m-CBT was delivered with high degree of adherence to treatment protocol. Number of completed sessions did not differ across conditions. **Was contamination acceptably low?** Yes. **Did either group receive additional interventions or have services provided in a different manner?** No. **Were outcomes relevant?** Yes **Were outcome measures reliable?** Yes. All outcome measures are recognised, validated instruments.	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Partly. Research procedure approved at University of Denver and community clinic review board. Consent provided by parents - unclear if adolescents also gave their own consent. **Were service users involved in the study?** No. Service users involved as participants only. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes.	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ Despite US context both m-CBT and usual care (client-centred or psychodynamic interventions) are similar to the UK context. **Overall validity rating:** - Key limitations of the study include relatively small sample size, no male participants completed the study. Unclear what the analysed sample size was, due to exclusion of male participants.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Was the allocation method followed? Yes. **Is blinding an issue in this study?** Blinding not possible. **Did participants reflect target group?** Yes. **Were all participants accounted for at study conclusion?** Yes. Seven participants were missing outcome data at 16 weeks - this group differed only from those for whom data was available in terms of number of sessions attended. Important to note: it appears that all 7 missing participants were male. Data were therefore analysed for females only.	**Were all outcome measurements complete?** Yes. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Partly. Final assessment conducted at 16 weeks - relatively short follow up time. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Yes. There were some differences in groups in terms of location of treatment, but this was included as a covariate. **Was intention to treat (ITT) analysis conducted?** Yes. Last observation carried forward for missing data. **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported.	Young people with depression and who had experienced abuse. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to Response. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **Does the study have a UK perspective?** No. US study.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Not reported, however relatively small sample size. **Were the estimates of effect size given or calculable?** No. **Were the analytical methods appropriate?** Yes. **Was the precision of intervention effects given or calculable? Were they meaningful?** Yes. **Do conclusions match findings?** Yes.		

10. Trowell J, Kolvin I, Weeramanthri T et al. (2002) Psychotherapy for sexually abused girls: psychopathological outcome findings and patterns of change. The British journal of psychiatry: the journal of mental science 180: 234–47

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Study aim 'To compare the relative efficacy of focused individual or group therapy in symptomatic sexually abused girls, and to monitor psychiatric symptoms for persistence or change' (p234). **Description of theoretical approach?** Partly. Study points out links between child sexual abuse and mental health problems. However theory	**Was the exposure to the intervention and comparison as intended?** Yes. **Was contamination acceptably low?** Yes. **Did either group receive additional interventions or have services provided in a different manner?** No.	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Partly. Ethical approval not mentioned. Both children and guardians had to give informed consent to participate. Study design meant that all	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + However, lack of reporting in relation to ethical approval is a concern. **Overall validity rating:** - Lack of clarity regarding statistical

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
base for the two treatment models not given. **How was selection bias minimised?** Randomised. **Was the allocation method followed?** Yes. **Is blinding an issue in this study?** Blinding not possible. Not possible to blind participants to condition (group versus individual therapy). Not reported whether assessors were blind to participant condition. **Did participants reflect target group?** Yes. **Were all participants accounted for at study conclusion?** Yes. Some missing data imputed using Last Observation Carried Forward.	**Were outcomes relevant?** Yes. **Were outcome measures reliable?** Yes. Standardised validated instruments used for all measures. **Were all outcome measurements complete?** Not reported. Some of the data relating to K-SADS do not seem to have been reported. **Were all important outcomes assessed?** Yes. **Were there similar follow-up times in exposure and comparison groups?** Yes. **Was follow-up time meaningful?** Yes. Two years. **Were exposure and comparison groups similar at baseline? If not, were these adjusted?** Partly. Baseline scores on outcome measures varied across groups. Authors have aimed to correct for this by conducting ANCOVA, but this is poorly reported.	children meeting the inclusion criteria received an intervention. **Were service users involved in the study?** No. Service users involved as participants only. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Girls who have experienced sexual abuse. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to response. **(For effectiveness questions) Are the study outcomes relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes.	analysis, including calculation of effect sizes, and unclear presentation of data are significant limitations in this study.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Was intention to treat (ITT) analysis conducted? Partly. Yes but missing data only imputed for ANCOVA calculations (not in Tables 2 and 3). **Was the study sufficiently powered to detect an intervention effect (if one exists)?** Not reported. However, relatively small sample size. **Were the estimates of effect size given or calculable?** Partly. Not all relevant effect sizes are reported (only those where d>0.5). When the reviewing team have attempted to calculate effect sizes from raw data in the paper, different estimates are obtained. It is therefore unclear what method the authors used to calculate effect size. **Were the analytical methods appropriate?** Partly. Analysis and presentation of findings in the paper is problematic: - Results tables show only changes in scores across measures, rather than raw data. In Table 3 it is states that 'high scores represent the most impairment', yet scores		

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	appear to increase for both groups at each time point, suggesting that impairment worsened throughout the intervention. Text elsewhere in the document suggests that this is a typo. - Effect sizes have only been reported where they exceed 0.5 - the reviewing team query the validity of this. - It is unclear whether standard deviations reported in the tables refer to the raw scores, or increase/decreases in scores - Potentially linked to the above, when the reviewing team has attempted to recalculate effect sizes based on data provided in the tables, different estimates are obtained - Reporting of 2-way ANCOVA reports only one F value (there should be two main effects and one interaction term). **Was the precision of intervention effects given or calculable? Were they meaningful?** No. **Do conclusions match findings?** Yes.		

Research question 16 – Findings tables

1. Barbe RP, Bridge AJ, Birmaher B et al. (2004) Lifetime history of sexual abuse, clinical presentation, and outcome in a clinical trial for adolescent depression. Journal of Clinical Psychiatry 65: 77–83

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The study aimed to determine what impact sexual abuse had on clinical depression and treatment outcome in depressed adolescents. **Methodology:** RCT. Participants were randomised to one of three treatment conditions. NB. Although participants were randomised to 1 of three treatment conditions this study only reports on those randomised to either cognitive-behavioural therapy or nondirective supportive therapy (n=72) due to the fact that only one participant assigned to systemic behavioural therapy was determined to have a history of sexual abuse. **Country:** USA.	**Participants:** Children and young people. Adolescents between the ages of 13 and 18 meeting criteria for DSM-III-R major depression with a score greater than or equal to 13 on the Beck Depression Inventory. Exclusion criteria appear to be; psychosis, bipolar disorder, substance abuse, ongoing abuse (physical or sexual), or obsessive-compulsive disorder. History of sexual abuse was determined by the question 'Have you ever been sexually abused by a relative, acquaintance, or stranger?' (p79). **Sample characteristics:** • Age – Mean age in years (SD) – Sexually abused group 15.7 (1.4). Never sexually abused group 15.9 (1.5). • Sex – Percentage female - Sexually abused group 90 %. Never sexually abused group 72.6%. • Ethnicity – Percentage White - Sexually abused group 60%. Never sexually abused group 83.9%. • Religion/belief – Not reported. • Disability – Not reported. • Long term health condition – Not reported.	**Statistical data - children and young people's health and wellbeing at treatment completion -** Rate of major depression - Participants randomised to the cognitive behavioural therapy group who had a history of sexual abuse had a higher rate of major depression at treatment completion than non-abused participants, but this was not significant 40% (2/5) vs. 13.3% (4/30), Fisher exact test p=.1912, effect size [phi]=0.248. For participants without a history of sexual abuse, membership of the intervention group was associated with significantly lower rates of major depression at end of treatment than for the control group (p=0.02). However, this effect was not observed for those with a history of sexual abuse. However, the difference in effect between sexually abused and non-sexually abused individuals was not significant (chi-square=0.64, df=1, p=0.43). **Narrative findings - children and young people's health and wellbeing at treatment completion –** Rate of DSM-III-R major depression – For participants without a history of sexual abuse, membership of the intervention group was associated with significantly lower rates of major depression at end of treatment than for the control group. However, this effect was not observed for those with a history of sexual abuse. The authors interpret this as meaning that a history of sexual abuse means that individuals are 'less likely to respond to treatment' (p81). However, the difference in effect between sexually abused	**Overall assessment of internal validity:** - A key limitation is the referral of participants not showing progress to 'open treatment'. Also, the authors note that randomisation included 'balancing' for number of parents in household, gender, suicidality. It is unclear whether this refers to a priori stratification, or a post hoc adjustment to randomisation, which would in turn negate the randomisation process. **Overall assessment of external validity:** ++ **Overall validity rating:** -

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Source of funding: Government – USA Institute for Mental Health.	• Sexual orientation - Not reported. • Socioeconomic position – Determined using the Hollingshead Four-Factor Index of Social Status (unclear how socio-economic status is linked to each grade) Percentage in each grade. Sexually abused group; I=10%; II=30%; III=20%; IV=30%; V=10%. Never sexually abused group; I=6.5%; II=11.3%; III=27.4%; IV=40.3%; V=14.5%. • Type of abuse - Sexual abuse was determined for 11 participants (cognitive behavioural therapy group n=6; systemic behavioural family therapy n=4; nondirective supportive therapy n=1). NB The analysis reported here is restricted to those participants assigned to either cognitive behavioural therapy or nondirective supportive therapy as only one participant assigned to systemic behavioural family therapy was determined to have experienced sexual abuse. The question used to determine this was a dichotomous yes/no question meaning that no further details on extent, frequency, severity, etc., of sexual abuse could be determined. • Percentage of participants with a history of physical abuse – Sexually abused group 20%. Never sexually abused group 29.3%.	and non-sexually abused individuals was not significant, meaning that the basis for the conclusion is weak.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** Sample size – 107 participants were randomised but analysis is restricted to the 72 participants assigned to either cognitive behavioural therapy (n=37) or non-directive supportive therapy (n=35) as only one participant assigned to systemic behavioural family therapy was determined to have experienced sexual abuse. • Comparison numbers – Cognitive behavioural therapy - n=37. Sexually abused n=6, never sexually abused n=31. Systemic behavioural family therapy - n=35. Sexually abused n=1, never sexually abused n=34. Nondirective supportive therapy - n=35. Sexually abused n=4, never sexually abused n=31. NB The analysis reported here is restricted to those participants assigned to either cognitive behavioural therapy or nondirective supportive therapy as only one participant assigned to systemic behav-		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	ioural family therapy was determined to have experienced sexual abuse. **Intervention:** Cognitive behavioural therapy. The authors do not provide any details on this intervention other than noting that it is derived from Beck et al. (1979). • Delivered by – Cognitive behavioural therapy – Not reported. • Delivered to - Cognitive behavioural therapy - Not reported. • Duration, frequency, intensity, etc. – Cognitive behavioural therapy – 12 to 16 sessions delivered over 12-16 weeks. • Key components and objectives of intervention – Cognitive behavioural therapy - Not reported. • Content/session titles - Cognitive behavioural therapy - Not reported. • Location/place of delivery - Cognitive behavioural therapy - Not reported. • Describe comparison intervention – • Nondirective supportive therapy – 12 to 16 sessions delivered over 12–16 weeks. The only other detail provided in relation to this intervention is that it is designed to '… control for the nonspecific effects of psychotherapy and consisted of the		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	provision of support, affect clarification, and active listening' (p78). • Systemic behavioural family therapy - 12 to 16 sessions delivered over 12–16 weeks. The only other detail provided in relation to this intervention is that it combines functional family therapy with a problem solving model created by Robin and Foster (1989). NB Participants assigned to this intervention were excluded from analysis as only one participant was determined to have experienced sexual abuse. **Outcomes measured:** • Children and young people's health and wellbeing - A number of measures of health and wellbeing were utilised in the study. However, only one was analysed in terms of the interaction between treatment group and abuse history – this was rate of DSM-III-R major depression at the end of treatment assessed using the interviewer Kiddie Schedule for Affective Disorders and Schizophrenia - Present Episode Version (Chambers et al. 1985) and the Kiddie Schedule for Affective Disorders and Schizophrenia – Epidemiologic Version (Overaschel et al. 1982). • Service outcomes - No service outcomes were analysed in terms of		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	the interaction between treatment group and abuse history. **Follow-up:** There were 7 follow-up assessments in total. These took place at the sixth treatment session; at treatment completion (between the 12th and 16th sessions); every 3 months during the first year and at 24 months post-treatment. NB. Only data collected at post-treatment and 24 months post-treatment are reported here and only the primary outcome of interviewer assessed rate of DSM-III-R major depression was measured at both time points. **Costs?** No. Cost/resource use information is not reported.		

2. Carpenter J, Jessiman T, Patsos D et al. (2016) Letting the future in: A therapeutic intervention for children affected by sexual abuse and their carers. London: NSPCC

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim Research questions: '1. What are the outcomes for children and young people affected by sexual abuse of providing Letting the Future In in NSPCC services? 2. What is the cost-effectiveness	**Participants** Children and young people. Children and young people aged 4 to 17 'affected by sexual abuse' (p10). Caregivers and families - 'Safe carers' (p10) of children and young people aged 4 to 17 affected by sexual abuse.	**Effect sizes** **Children and young people's health and wellbeing** A - Primary outcomes (clinical or significant difficulty level scores on TSCC or TSCYC) 1. Older children and young people (8 and over) Results are reported for 'analysis completers' (those for whom T1 and T2 data were available) and 'intention to treat' in which missing data have been imputed statistically (using demographic and other variables	**Overall assessment of internal validity** + **Overall assessment of external validity** ++ **Overall validity score** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
of this service? 3. What is the effectiveness of the support intervention received by the 'safe' carers?' (p10). **Methodology** RCT including cluster. Intervention and wait-list control. **Country** UK. **Source of funding** Voluntary/charity - NSPCC Other commercial source - Impetus - The Private Equity Foundation.	**Sample characteristics:** • Age – Children: Mean age 10.7, age range 3 (note exec summary states 4) to 17 Carers: 20–29 12%, 30-39 36%, 40-49 27%, 50+ 9%, unknown 18%. • Sex – Children: 26% male, 74% female Carers: Male 8%, female 89%, unknown 4%. • Ethnicity – Children: Black and minority ethnic 9% Carers: Black and minority ethnic 4%. • Religion/belief – Not reported. • Disability – Children: One or more disabilities 17% Carers: Not reported. • Long term health condition – Not reported. • Sexual orientation – Not reported. • Socioeconomic position – Not reported. • Type of abuse – Sexual abuse Age at onset of abuse: Less than 3 years old 20%, 3-7 38%, 8-12 26%, 13+ 16% Nature of abuse: Non-contact 12%, online abuse 4%, inappropriate touching 66%, penetration or attempted penetration 49%, sexual abuse with violence 5%. Number of known incidents of abuse: One 28%, 2-4 21%, 5+ 38%, unknown 14%. Duration between onset and discovery of sexual abuse: Less than 6 months	likely to affect treatment outcomes). The intention to treat data are reported in the narrative summary as these provide the more conservative estimates. Separate analyses have been completed for: - Clinical level scores only - Clinical level and 'significant difficulty' scores 1.1 Clinical level scores on one more TSCC subscales Results have been analysed using McNemar's test for the difference between correlated proportions. However, test statistics are not reported, only p values. 1.1.1 Analysis completers There was a significant reduction in the proportion of children and young people with clinical level scores on one or more TSCC subscales in the intervention condition (p=0.029) but not the waitlist condition (p=1.00). 1.1.2 Intention to treat Using the ITT approach, the reduction in the proportion of children and young people with clinical level scores on one or more TSCC subscales in the intervention condition was only marginally significant (p=0.065). Wait list condition was non-significant (p=0.839). 1.2 Combined clinical and significant difficulty scores 1.2.1 Analysis completers There was a significant reduction in the proportion of children and young people with clinical level or significant difficulty scores on 1 or more TSCC subscales in	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	52%, 6-12 months 20%, more than 12 months 28%. Relationship of perpetrator with child: Intrafamilial 65%, extrafamilial 35%. Number of perpetrators: Single perpetrator 80%, 2+ perpetrators 10%, unknown 10%. Perpetrator gender: Male 93%, female 5%, male and female 2%. Perpetrator age: Adult 58%, young person over 14 years 22%, young person 11-13 years 9%, children aged 10 and under 11%, unknown 3%. Looked after or adopted status – 12% were looked after. • Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size:** Comparison numbers - Children: 114 Carers: 76 Intervention numbers - Children: 128 Carers: 89 Sample size - Children=242 Carers=165 Describe intervention Therapeutic intervention which is largely psychodynamic and 'grounded in an understanding of trauma, attachment and resilience' (p10), and draws on methods such as counselling and socio-educative approaches.	the intervention condition (p=0.001) but not the wait-list condition (p=0.581). 1.2.2 Intention to treat There was a significant reduction in the proportion of children and young people with clinical level or significant difficulty scores on 1 or more TSCC subscales in the intervention condition (p=0.016) but not the wait-list condition (p=1.00). 1.3 TSCC subscale analysis 1.3.1 Analysis completers NOTE - some very small cell sizes for this analysis. The study found that those in the intervention group, but not the control group showed significant shift away from clinical/significant difficulty level scores on the following subscales: anxiety (intervention p=0.035, control p=NS), post-traumatic stress (p=0.011, p=0.118), and dissociation (p=0.043, p=0.629). For 1 item (sexual concerns - general) the wait list group showed a significant improvement (p=0.003) but the intervention group did not. All other subscales did not show significant differences for either group (depression, anger, dissociation - overt, dissociation - fantasy, sexual concerns - preoccupation, sexual concerns - distress. 1.3.2 Intention to treat Analysis of the TSCC subscales using imputed data provided a different picture. In this case, the intervention group showed significant improvement in relation to anger (intervention p=0.003, control p=0.076), post-traumatic stress (p=0.020, p=0.108) and dissociation - general (p=0.020, p=0.108). For 1 item (sexual	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	The intervention emphasises the therapeutic 'attunement of the practitioner to the child's emotional responses' (p10). Delivered by Therapists, also described as 'children's services practitioners (CSPs)... [with] varying levels of experience and training' (p21). Unclear if all were qualified psychotherapists. Delivered to Children aged 4 to 17 affected by sexual abuse and their non-abusing carer. Duration, frequency, intensity, etc. Intervention delivered over 6 months. Children offered up to four therapeutic assessment sessions and 20 intervention sessions, extended to 30 as necessary. Median number of sessions was 18.4 for children aged 8 and over (14.6 were individual work with child, 2.5 safe carer work, 1.4 safe carer and joint child work). Median number of sessions was 19.6 for children under 8 (13.2 individual work with child, safe carer work 3.4, safe carer and child joint work 3.0). Safe carers are offered up to 8 sessions. The median number of sessions was 6.9. Key components and objectives of intervention The framework for intervention is	concerns - general) the wait list group showed a significant improvement (p=0.008) but the intervention group did not (p=0.115). All other subscales did not show significant differences for either group (anxiety, depression, dissociation-overt, dissociation - fantasy, sexual concerns - preoccupation, sexual concerns - distress). 2. Young children (under 8) 2.1 Clinical level scores on one more TSCC subscales 2.1.1 Analysis completers Neither the intervention nor control groups showed a significant reduction in the proportion of children with clinical level scores on 1 or more TSCC subscales (intervention p=0.687; control p=1.000). 2.1.2 Intention to treat Results not reported. 2.2 Combined clinical and significant difficulty scores 2.2.1 Analysis completers Results not reported. 2.2.2 Intention to treat Neither the intervention nor control groups showed a significant reduction in the proportion of children with clinical level or 'significant difficulty' level scores on 1 or more TSCC subscales (intervention p=0.625; control p=1.000). 2.3 TSCYC subscale analysis. 2.3.1 Analysis completers.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	based on Bannister's (2003) Recovery and Regeneration Model, which is largely psychodynamic and draws on psychodrama (Moreno 1983), play therapy (Gil 1991) and attachment theory (Bowlby 1969). Bannister's model uses three key phases: 1. Assessment of the child's developmental needs and how the process of development may have been interrupted by sexual abuse 2. Action phase - building a positive relationship with the worker including developing boundaries and 'confirming the child's feelings and identity' (p21). This draws on techniques such as play, art, drama and stories. 3. Resolution phase - child is encouraged to express feelings and develop self-awareness and relationships. Despite the influence of Bannister's framework the authors describe Letting the Future In as 'deliberately multi-theoretical', and state that a number of constructs have been used to build: - a value base - an underpinning knowledge base including knowledge of child development, attachment theory and resilience factors - a skills base including using the therapeutic relationship and motivational interviewing and elements of trauma-focused CBT (psycho-educative work, development of parenting skills, joint carer-child sessions). In terms of the	The intervention groups showed significant reductions in two subscales: post-traumatic stress - intrusion (intervention p=0.022; control p=0.039) and post-traumatic stress - avoidance (p=0.039, p=0.375). There were no significant reductions on the remaining seven subscales: anxiety, depression, anger, post-traumatic stress - arousal, post-traumatic stress - total, dissociation or sexual concerns. 2.3.2 Intention to treat Using the intent to treat data there were no significant improvements in any of the TSCYC subscales for either intervention or control group. The control group deteriorated in relation to the depression (p=0.022) and sexual concerns (p=0.039) subscales. B - Secondary outcomes Secondary outcomes were the changes in mean scores on the TSCC and TSCYC subscales, analysed using ANCOVA. There were no statistically significant differences on any subscales for either older or younger children (no data reported). C - Effectiveness analysis The study had planned to examine whether children who received four or more sessions did better than those who dropped out early. However, 87% of participants received 4 or more sessions, meaning that this comparison was not possible. D - Follow on one year from baseline 1. Older children 1.1 Proportion of children reporting clinically significant scores on one or more TSCC subscales 1.1.1 Analysis completers	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	carer component of the intervention, the key aims are to help carers to: 'process the impact of discovering the child's sexual abuse; create a social environment that facilitates their children's recovery; provide emotional warmth alongside structure and routine; help their child feel safe, and collaborate in the process of their child in re-authoring their trauma narrative' (p22). The intervention is guided by a 2-stage assessment: 1. Referral assessment - to gather initial information and judge suitability of referral 2. Assessment of therapeutic need, based on child's self-completion of the Trauma Symptoms Checklist (TSCC) or Trauma Symptoms Checklist for Young Children (TSCYC) on the child's behalf (for children aged under 8). Content/session titles Children: Records were kept of primary content of individual sessions with each child. Content of sessions as follows: Creative therapies (20% of sessions) Awareness and management of feelings (17%) Counselling (15%) Identity and self-esteem (9%) Socio-educative (9%) Symbolic play (8%) Solution-focused brief therapy (7%) Agreement and boundary formation (5%) Using scales and tools (4%) Attachment based (3%) Gradual exposure (1%) Trauma-focused CBT	For the children for whom data were available at all three time points (n=34) the following pattern of results was observed: at T1 52.9% of young people had clinically significant scores, T2 23.5%, T3 44.1%. Cochran's Q test was carried out and was significant (Q=8.316, p=0.16), however it is unclear whether all differences are significant - no post hoc testing undertaken. 1.1.2 Intent to treat analysis The study reports that the intervention group showed a statistically significant improvement in the proportion of children reporting clinical level scores between T1 and T2 (p=0.041). There was a non-significant increase in children with clinical or significant difficulty level scores between T2 and T3 (p=0.263). 1.2 Proportion of children reporting clinically significant or significant difficulty scores on 1 or more TSCC subscales 1.2.1 Analysis completers Not reported. 1.2.2 Intent to treat A similar pattern was observed, of a significant decrease in the proportion of children with clinical or significant difficulty level scores between T1 and T2 (p=0.020) and a non-significant increase between T2 and T3 (p=0.503). 2. Younger children 2.1 Proportion of children reporting clinically significant scores on one or more TSCYC subscales 2.1.1 Analysis completers	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	(1%) Carers: 'Generally based on counselling and awareness and management of feelings, together with socio-educative work' (p62). The following topic areas are covered: 'Helping the carer express and process the personal impact of discovering that their child was sexually abused; educating carers about the nature and consequences of sexual abuse; helping carers consider how they can support their child, assessing the safe carer's capacity for joint sessions and preparing for joint sessions with the child/young person and safe carer' (p24). Location/place of delivery Not reported. Describe comparison intervention Waiting list. **Outcomes measured** Children and young people's health and wellbeing outcomes Change in proportion of children with clinical levels of symptoms or problematic behaviour or significant difficulties, as assessed using the Trauma Symptoms Checklist (TSCC) (Briere 1996) for children aged 8 and over, and Trauma Symptoms Checklist for Young Children (TSCYC) (Briere et al. 2001) for children aged under 8. Caregiver/parent health and wellbeing outcomes	For children for whom data were available at all 3 time points (n=15) there was a non-significant decrease in the proportions of children with clinically significant scores on one or more subscales (p=0.687) but a marginally significant decrease between T2 and T3 (p=0.063). The authors suggest this may be because younger children tended to receive more sessions in between T2 and T3 than older children. 2.1.2 Intent to treat Not reported. 2.2. Proportion of children reporting clinically significant or significant difficulty scores on 1 or more TSCC subscales 2.2.1 Analysis completers Not reported. 2.2.2 Intent to treat Not reported. **Caregiver health and wellbeing** 1. Proportion of carers with Parenting Stress scores at or above clinical threshold Data are reported for analysis completers only - no intent to treat analysis has been conducted. The study found that there was little change in scores for either intervention or control between T1 and T2. At 1 year follow up (T3) significant improvements in total stress were observed for both the intervention (p=0.016) and control groups (p=0.021). Overall, the change in proportion of clinical level scores in the intervention group was significant (Q=7.2, p=0.027).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Change in parental stress as measured by the Parenting Stress Index (Abidin 1995).		

3. Danielson CK, McCart MR, Walsh K, et al. (2012). Reducing substance use risk and mental health problems among sexually assaulted adolescents: A pilot randomized controlled trial. Journal of Family Psychology 26: 628–35

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim To evaluate the differential efficacy of Reduction through Family Therapy (RRFT) and treatment as usual (TAU) in reducing substance use (SU) problems (including early initiation), mental health symptoms, and risky sexual behavior among adolescent CSA victims. **Methodology** RCT including cluster. **Country** Not UK. USA - Charleston, South Carolina **Source of funding** Pharmaceutical -	**Participants** Children and young people - Thirty adolescents (aged 13–17 years; M=14.80; SD=1.51) who had experienced at least 1 sexual assault and their caregivers were randomized to RRFT or treatment as usual (TAU) conditions. Caregivers and families - Caregivers: 72.6% biological parents, 17.1% other family members, 10.3% non-familial guardians. **Sample characteristics:** • Age – Adolescents aged 13–17 years (M=14.80; SD=1.51). • Sex – Approximately 88% of the sample was female. • Ethnicity – 46% were African American (37.5% white, 4.2% Native American, 8.3% bi-racial, 4% Hispanic). Religion/belief – Not reported. • Disability – Not reported.	**Effect sizes** **Children and young people's health and wellbeing** FROM TABLE 2: Mixed-Effect Regression Models for Treatment Outcome Measures UCLA PTSD-Adolescent Intercept: RRFT β37.24 SE: 4.02 df=28 p=<.001 (29.36, 45.12) TAU v. RRFT β−4.60 SE: 5.65 df=28 p=0.422 (−15.65, 6.45) Slope: RRFT β−1.19 SE: 0.24 df=81 p=<.001 (−1.66, −0.72) TAU v. RRFT β0.42 SE: 0.34 df=81 p=0.215 (−0.25, 1.09) UCLA PTSD-Parent Intercept: RRFT β39.09 SE: 3.66 df=28 p=<.001 (31.92, 46.26) TAU v. RRFT β−12.97 SE: 5.14 df=28 p=0.018 (−23.04, −2.90)	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** + **Overall validity rating:** + + (moderate) external validity (related to narrow match of study outcomes relative to review scope) but ++ internal validity = + moderate overall validity.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
The study was supported by grant award K23DA018686 from the National Institute on Drug Abuse (NIDA; PI: Danielson) and a Young Investigator Award from NARSAD (PI: Danielson).	• Long term health condition – Not reported. • Sexual orientation – Not reported. • Socioeconomic position – Not reported. • Type of abuse – Sexual: experienced at least 1 lifetime CSA that could be recollected by the youth (defined as unwanted/forced vaginal or anal penetration by an object, finger, or penis; oral sex; or touching of one's genitalia). Mean time since most recent assault was 3.7 years (SD=3.8). • Looked after or adopted status – Not reported. • Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size** Comparison numbers - 15 in the TAU group. Intervention numbers - 15 in the RRFT group. Sample size - Participants included 30 treatment seeking adolescent CSA victims. 15 were included in the treatment group (RRFT) and 15 in the 'control' group (TAU). **Intervention** **Describe intervention**	Slope: RRFT β−1.46 SE: 0.21 df=82 p=<.001 (−1.87, −1.05) TAU v. RRFT β0.87 SE: 0.29 df=82 p=0.004 (0.30, 1.44) Child Depression Inventory Intercept: RRFT β60.42 SE: 2.73 df=28 p=<.001 (55.07, 65.77) TAU v. RRFT β−8.54 SE: 3.83 df=28 p=0.034 (−16.05, −1.03) Slope: RRFT β−0.87 SE: 0.17 df=81 p=<.001 (−1.20, −0.54) TAU v. RRFT β0.52 SE: 0.24 df=81 p=0.036 (0.05, 0.99) Behavioural Assessment System for Children - Internalising Intercept: RRFT β67.45 SE: 2.47 df=28 p=<.001 (62.61, 72.29) TAU v. RRFT β−10.71 SE: 3.44 df=28 p=0.004 (−17.45, −3.97) Slope: RRFT β−1.06 SE: 0.14 df=81 p=<.001 (−1.33, −0.79) TAU v. RRFT β0.53 SE: 0.20 df=81 p=0.008 (0.14, 0.92) Behavioural Assessment System for Children - Externalizing Intercept: RRFT β66.94 SE: 3.27 df=28 p=<.001 (60.53, 73.35) TAU v. RRFT β−6.43 SE: 4.52 df=28 p=0.166 (−15.29, 2.43) Slope: RRFT β−0.90 SE: 0.22 df=81 p=<.001 (−1.32, −0.46)	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	In an effort to better address the multiple clinical problems commonly experienced by victims of CSA, the authors recently developed an integrated treatment protocol called Risk Reduction through Family Therapy (RRFT; Danielson et al. 2010a). RRFT builds upon the principles and interventions applied in empirically-supported treatments for adolescent SU (Multisystemic Therapy/MST; Henggeler et al. 2002), PTSD and depression (TF-CBT), and other negative sequelae (e.g., risky sexual behaviors; DiClemente et al., 2004). Importantly, the model utilises exposure-based techniques (i.e., where one learns to gain control of the distress induced by trauma-related cues) from TF-CBT to address youths' trauma symptoms, as well as involvement of the family, which has been demonstrated to be beneficial in the treatment of youth PTSD (Gilboa-Schechtman et al. 2010) and SU problems (Henggeler et al. 2002). Data from an open pilot suggest that RRFT is feasible and potentially efficacious in reducing risk for SU problems and decreasing PTSD and depression among youth (Danielson et al., 2010a). RRFT—RRFT was developed to reduce risk of SU and other high-risk behaviors and trauma-related psychopathology in adolescents	TAU v. RRFT β0.42 SE: 0.31 df=81 p=0.181 (−0.19, 1.03) Time Line Follow Back Days w. Use Intercept: RRFT β0.21 SE: 0.70 df=28 p=0.764 (−1.16, 1.58) TAU v. RRFT β−2.17 SE: 1.01 df=28 p=0.040 (−4.13, −0.21) Slope: RRFT β−0.17 SE: 0.01 df=81 p=<.001 (−0.19, −0.15) TAU v. RRFT β 0.30 0.03 81 <.001 (0.24, 0.36) Any Sexual Partners Intercept: RRFT β −0.04 0.56 27 0.948 (−1.14, 1.06) TAU v. RRFT β 0.37 0.76 27 0.633 (−1.12, 1.86) Slope: RRFT β −0.12 0.05 82 0.026 (−0.22, −0.02) TAU v. RRFT β −0.01 0.07 82 0.912 (−0.15, 0.13) Family Environment Scale Cohesion-Adolescent Intercept: RRFT β 40.16 3.78 28 <.001 (32.75, 47.57) TAU v. RRFT β 12.80 5.31 28 0.023 (2.39, 23.21) Slope: RRFT 0.76 0.21 81 0.001 (0.35, 1.17) TAU v. RRFT β −1.03 0.30 81 0.001 (−1.62, −0.44) Family Environment Scale Cohesion-Parent Intercept: RRFT β 45.25 2.74 28 <.001 (39.88, 50.62) TAU v. RRFT β 12.27 3.81 28 0.003 (4.80, 19.74) Slope: RRFT β 0.85 0.21 82 <.001 (0.44, 1.26) TAU v. RRFT β −0.79 SE: 0.30 df=82 p=0.010 (−1.38, −0.20) Family Environment Scale Conflict-A Intercept: RRFT β 58.69 SE: 2.72 df=28 p=<.001 (53.36, 64.02)	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	who have experienced CSA. RRFT integrates the nine guiding principles of MST with other empirically-supported interventions with similar theoretical rationales targeting similar populations, including TF-CBT and psychoeducation strategies for prevention of high-risk sexual behavior (e.g., DiClemente et al. 2004) and sexual revictimization (Marx et al. 2001). The RRFT protocol is devised into 7 treatment components: Psychoeducation, Coping, Family Communication, Substance Abuse, PTSD, Healthy Dating and Sexual Decision Making, and Revictimisation Risk Reduction and is administered through weekly, 60–90 minute sessions with adolescents and caregivers (meeting individually with the therapist and as a family). RRFT utilizes a clinical pathways approach that is driven by youth symptomatology (i.e., the order of and time spent on each component is determined by the needs of each youth and family). Several theories are incorporated into RRFT via multiple intervention strategies. First, RRFT is guided by ecological theory, which proposes that an adolescent's behavior is influenced by multiple social and environmental contexts, including the family, peer network, school, and community (Bronfenbren-	TAU v. RRFT β −9.15 SE: 3.80 df=28 p=0.023 (−16.60, −1.70) Slope: RRFT β −1.13 SE: 0.19 df=81 p=<.001 (−1.50, −0.76) TAU v. RRFT β 0.92 SE: 0.27 df=81 p=0.001 (0.39, 1.45) Family Environment Scale Conflict-P Intercept: RRFT β 54.70 SE: 2.72 df=28 p=<.001 (49.36, 60.02) TAU v. RRFT β −6.44 SE: 3.81 df=28 p=0.102 (−13.91, 1.03) Slope: RRFT β −0.79 SE: 0.17 df=82 p=<.001 (−1.12, −0.46) TAU v. RRFT β 0.47 SE: 0.25 df=82 p=0.058 (−0.02, 0.96)	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	ner 1979). RRFT, like other ecological models (e.g., MST), adopts a family-based approach to intervention and encourages therapists to intervene in multiple social systems. As part of this ecological model, Strategic Family Therapy (e.g., Haley 1976) is utilised to help the family define problems (in behaviourally specific terms – and from the perspective of both of the adolescent and the caregiver) and work together to solve those problems. The family is involved across all 7 treatment components. Mowrer's Two-Factor Theory (Mowrer, 1960) and negative reinforcement theory (Baker et al. 2004) also are applied in targeting PTSD, SU, and their overlap in RRFT. Additional information about the RRFT model can be found in Danielson et al. (2010a). **Delivered by** Clinical psychology graduate students. Therapy was delivered through both an outpatient clinic and an outreach programme offered at the same clinic. **Delivered to** Adolescents visiting a weekly mental health clinic.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Duration, frequency, intensity, etc. Mean treatment length for RRFT was 23 sessions (SD=13); mean number of weeks in treatment was 34 (SD=17). Treatment completers were defined as having completed 5 of 7 RRFT components. Key components and objectives of intervention Location/place of delivery The trial was conducted in an urban clinic specialising in the treatment of trauma. Therapy was delivered through both an outpatient clinic and an outreach program offered at the same clinic. Describe comparison intervention Treatment as Usual (TAU) – range of treatments. Clinic chart reviews for those in the TAU condition showed that no one treatment was consistently delivered across youth and families assigned to this condition. **Recognition tool** Describe recognition tool Mental Health Symptoms – PTSD symptoms were assessed with the UCLA PTSD Index for DSM-IV-Adolescent & Caregiver versions (Steinberg et al. 2004). The Child Depression Inventory (CDI; Kovacs 1983) was used to assess depression symptoms and the Behavioral Assessment System for Children		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	(BASC-2; Reynolds & Kamphaus 1992) (parent and youth self-report) measured participants' internalizing and externalising symptoms. Internal consistency for all measures was high in the current sample (Cronbach's alphas >.84). Substance Use and Substance Use Risk Factors—The Time Line Follow Back Interview (TLFB), a well-established method of assessing SU (Sobell & Sobell 1996), was conducted with each participant to identify specific amounts of alcohol and drugs consumed over the past 90 days. Urine drug screens were collected to validate self-reported SU (i.e., verify that denial of illicit drug use was accurate). The Cohesion and Conflict subscales of the Family Environment Scale (FES; Moos & Moos 1986) were completed by adolescents and caregivers, as these aspects of family environment have been linked with SU risk. The reliabilities of these subscales have been established (Boyd et al., 199; Cole & McPherson, 1993).Risky Sexual Behavior - Two items were used to assess: 1) number of consensual sexual intercourse partners over the past 3 months (including new sexual partners), and 2) whether or not the youth had been diagnosed with a sexually transmitted disease in the past 3 months.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Describe comparison tool/usual practice Same as intervention group. Participants completed measures of substance use, substance use risk factors (e.g., family functioning), mental health problems (i.e., posttraumatic stress disorder, depression, and general internalizing/externalising symptoms) and risky sexual behavior at four time points (baseline, post-treatment, and 3- and 6-month follow-up). **Outcomes measured** Children and young people's health and wellbeing outcomes Mental Health Symptoms PTSD symptoms measured using PTSD Index for DSM-IV-Adolescent & Caregiver versions (Steinberg et al. 2004). Depression measured using the Child Depression Inventory (CDI; Kovacs 1983). Internalising and externalising symptoms measured using Behavioral Assessment System for Children (BASC-2; Reynolds & Kamphaus, 192) (parent and youth self-report) Internal consistency for all measures was high in the current sample (Cronbach's alphas >.84). Substance Use and Substance Use Risk Factors Alcohol and drug use in previous 90 days measured using the Time Line Follow Back Interview (TLFB) Family		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	environment assessed using the Cohesion and Conflict subscales of the Family Environment Scale (FES; Moos & Moos 1986) completed by adolescents and caregivers, as these aspects of family environment have been linked with SU risk. Risky Sexual Behavior Number of consensual sexual intercourse partners over the past 3 months (including new sexual partners) Whether or not the youth had been diagnosed with a sexually transmitted disease in the past 3 months. **Recognition indicators measured** Trauma - Mental Health Symptoms - PTSD symptoms were assessed with the UCLA PTSD Index for DSM-IV-Adolescent & Caregiver versions (Steinberg et al. 2004). The Child Depression Inventory (CDI; Kovacs 1983) was used to assess depression symptoms and the Behavioral Assessment System for Children (BASC-2; Reynolds & Kamphaus 1992) (parent and youth self-report) measured participants' internalizing and externalising symptoms. Internal consistency for all measures was high in the current sample (Cronbach's alphas > .84). Substance abuse - Substance Use and Substance Use Risk Factors - The Time Line Follow		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Back Interview (TLFB), a well-established method of assessing SU (Sobell & Sobell 1996), was conducted with each participant to identify specific amounts of alcohol and drugs consumed over the past 90 days. Urine drug screens were collected to validate self-reported SU (i.e., verify that denial of illicit drug use was accurate). The Cohesion and Conflict subscales of the Family Environment Scale (FES; Moos & Moos 1986) were completed by adolescents and caregivers, as these aspects of family environment have been linked with SU risk. The reliabilities of these subscales have been established (Boyd et al. 1997; Cole & McPherson 1993). Risk factors - Risky Sexual Behavior - Two items were used to assess: 1) number of consensual sexual intercourse partners over the past 3 months (including new sexual partners), and 2) whether or not the youth had been diagnosed with a sexually transmitted disease in the past 3 months.		

4. Foa EB, McLean CP, Capaldi S (2013) Prolonged exposure vs supportive counseling for sexual abuse-related PTSD in adolescent girls: A randomized clinical trial. Journal of the American Medical Association 310: 2650–7

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The study aims to evaluate the	**Participants:** Children and young people. Female adolescents with a	**Statistical data - Baseline to post-treatment –**	**Overall assessment of internal validity:** ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
impact of counsellor-delivered prolonged exposure therapy in comparison to supportive counselling for adolescent girls with sexual abuse related post-traumatic stress disorder. **Methodology:** RCT. **Country:** USA – Philadelphia. **Source of funding:** Government - National Institute of Mental Health (R01 MH074505).	primary diagnosis of chronic or sub-threshold post-traumatic stress disorder resulting from experience of sexual abuse at least 3 months before assessment for study eligibility. Exclusion criteria were '…current suicidal ideation with intent, uncontrolled bipolar disorder, schizophrenia, conduct disorder, pervasive developmental disorder, initiation of psychotropic medication within the previous 12 weeks, and current inpatient psychiatric treatment. Adolescents with substance use or suicidality without imminent threat were not excluded' (p2651). Six participants with late stage pregnancies were also excluded. **Sample characteristics:** • Age – All participants were between the ages of 13 and 18. Mean age of total sample=15.3 years (15.0-15.7, 95% CI); mean age of intervention group=15.4 (14.9-15.8, 95% CI); mean age of comparison group=15.3 (14.7-15.9, 95% CI). • Sex – All participants were female. • Ethnicity – Total sample – Black n=34, White n=11, Hispanic n=10, Biracial n=2, 'other or no response' n=4; intervention group – Black n=19, White n=5, Hispanic n=3, Biracial n=0, 'other or no response'	Interviewer-rated post-traumatic stress disorder symptom severity (Child PTSD Symptom Scale–Interview) - Intervention: Significant improvement in symptom severity with large effect size; change in score=20.1; 95% CI 16.7-23.6; t57=11.35; p<.001; d=2.72. Comparison: Significant improvement in symptom severity with a large effect size; change in score 12.6; 95% CI 9.0-16.2; t60=6.83; p<.001; d=1.71. Between group differences: The intervention group showed significantly greater improvement in symptom severity, with a large effect size: Between-treatment difference in mean score=7.5; 95% CI 2.5-12.5; t59=2.93; p<.001; d=1.01. Rates of loss of post-traumatic stress disorder diagnosis (post-traumatic stress disorder module of the DSM-IV Schedule for Affective Disorders and Schizophrenia for School-Age Children) – Intervention: Significant decrease in rates of diagnosis, effect size not reported; Rate of loss of diagnosis: 83.3%; 95% CI 77.2%-85.4%; t270=8.92; p<.001. Comparison: Significant decrease in rates of diagnosis, effect size not reported; Rate of loss of diagnosis: 54%; 95% CI 49.1%-56.5%; t270=8.33; p<.001. Between group differences: The intervention group showed a significantly greater decrease in rate of diagnosis, effect size not reported; Difference in loss of diagnosis: 29.3%; 95% CI 20.2%-41.2%; t270=2.65; p<.01. Self-reported post-traumatic stress disorder symptom severity (Child PTSD Symptom Scale – Interview) -	**Overall assessment of external validity:** ++ **Overall validity rating:** ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	n=4; comparison group – Black n=15, White n=6, Hispanic n=7, Bi-racial n=2, 'other or no response' n=0. • Religion/belief – Not reported. • Disability – Not reported. • Long term health condition – Not reported. • Sexual orientation – Not reported. • Socioeconomic position – Not reported. • Type of abuse – All participants had experienced sexual abuse at least three months before assessment for study eligibility. No further details on this are provided, such as the period over which this occurred, perpetrator, etc.; however treatment and study participation were provided through a rape crisis centre and the authors note that non-offending primary guardians provided assent suggesting that the perpetrator in some instances may have been a family member or parent. • Looked after or adopted status – Not reported. • Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size:** • Comparison numbers – n=31.	Intervention: Significant improvement in symptom severity, effect size not reported; Change in score: 20.6; 95% CI 16.7-24.5; t202=10.23; p<.001. Comparison: Significant improvement in symptom severity, effect size not reported; Change in score: 14.4; 95% CI 10.3-18.5; t208=6.79; p<.001. Between group differences: The intervention group showed a significantly greater improvement in symptom severity, effect size not reported; Difference in changes in score: 6.2; 95% CI 1.2-11.2; t209=2.41; p=.02. Self-reported depression severity (Children's Depression Inventory) – Intervention: Significant improvement in depressive symptoms, effect size not reported; Change in score: 11.4; 95% CI 9.2-13.6; t135=10.4; p<.001. Comparison: Significant improvement in depressive symptoms, effect size not reported; Change in score: 6.5; 95% CI 4.0-9.0; t139=5.17; p<.001. Between group differences: The intervention group showed a significantly greater improvement in depressive symptoms, effect size not reported; Difference in changes in score: 4.9; 95% CI 1.6-8.2; t137=2.91; p=.008. Functioning (Children's Global Assessment Scale) - Intervention: Significant improvement in functioning, effect size not reported; Change in score: 17.8; 95% CI 13.2-22.4; t55=7.53; p<.001. Comparison: Significant improvement in functioning, effect size not reported; Change in score: 7.7; 95% CI 2.9-12.5; t58=3.13; p<.003.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Intervention numbers – n=30. • Sample size – n=61. **Intervention:** • Intervention category – Other - Prolonged exposure therapy. The authors note that the intervention is a modified version of that used with adults but do not provide details on this. • Describe intervention – The intervention comprises eight modules delivered in up to 14 weekly 60 to 90 minute sessions. Treatment includes 'in vivo' exposure (confronting trauma reminders in real life) and 'imaginal exposure' (revisiting and recounting the traumatic memory) (p2652). It is unclear, in the context of sexual abuse, what either in vivo or imaginal exposure comprise. • Delivered by – Masters level counsellors working at a rape crisis centre who usually provide supportive counselling. Counsellors attended a four day training workshop and group supervision was provided by 2 of the authors every other week. • Delivered to – Female adolescents between the ages of 13 and 18 with recent experience of sexual abuse.	Between group differences: The intervention group showed a significantly greater improvement in functioning, effect size not reported; Difference in changes in score: 10.1; 95% CI 3.4-16.8; t57=2.95; p=.008. **Post-treatment to 12 month follow-up –** Interviewer-rated post-traumatic stress disorder symptom severity (Child PTSD Symptom Scale–Interview) – Intervention: No significant improvement in symptom severity; p>.88. (No further statistical data presented). Comparison: No significant improvement in symptom severity; p>.88. (No further statistical data presented). Between group differences: No significant difference between groups in improvement of symptom severity p>.89. (No further statistical data presented). Rates of loss of post-traumatic stress disorder diagnosis (post-traumatic stress disorder module of the DSM-IV Schedule for Affective Disorders and Schizophrenia for School-Age Children) – Intervention: No significant decrease in rates of diagnosis; p>.19. (No further statistical data presented). Comparison: No significant decrease in rates of diagnosis; p>.19. (No further statistical data presented). Between group differences: No significant difference between groups in decrease in rates of diagnosis p > .57. (No further statistical data presented). Self-reported post-traumatic stress disorder symptom severity (Child PTSD Symptom Scale–Interview) - Intervention: No significant improvement in symptom severity; p>.19. (No further statistical data presented). Comparison: No significant improvement in symptom severity; p>.19. (No further statistical data presented).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Duration, frequency, intensity, etc. – Up to 14 weekly sessions between 60 and 90 minutes in duration. Participants also complete 'homework' between sessions. Prior to randomisation both groups participated in preparatory sessions (up to 3) which focused on case management concerns such as level of parental involvement, desire for treatment, legal processes and safety concerns (participants assessed as having active suicidal plans were excluded from the study). • Key components and objectives of intervention – The programme includes 8 modules which can be presented across a number of sessions according to the needs and abilities of each participant. No further details are provided such as the theories of change. • Content/session titles - • 1. Treatment rationale. • 2. Identification of index trauma, collection of information and breathing retraining exercises. • 3. Reactions to trauma. • 4. Rationale for in vivo exposure and '… confronting trauma reminders in real life …' (p2652), creation of an in vivo hierarchy, setting the in vivo homework.	Between group differences: No significant difference between groups in improvement of symptom severity p>.57. (No further statistical data presented). Self-reported depression severity (Children's Depression Inventory) – Intervention: No significant improvement in symptom severity; p > .19. (No further statistical data presented). Comparison: No significant improvement in symptom severity; p > .19. (No further statistical data presented). Between group differences: No significant difference between groups in improvement of symptom severity p > .57. (No further statistical data presented). Functioning (Children's Global Assessment Scale) - Intervention: No significant improvement in functioning; p > .19. (No further statistical data presented). Comparison: No significant improvement in functioning; p > .19. (No further statistical data presented). Between group differences: No significant difference between groups in functioning; p > .57. (No further statistical data presented). **Baseline to 12 month follow-up -** Interviewer-rated post-traumatic stress disorder symptom severity (Child PTSD Symptom Scale–Interview) – Intervention: Significant improvement in symptom severity with a large effect size; Change in score: 19.8; 95% CI 16.8-22.8; t52 = 12.87; p < .001; d = 2.67. Comparison: Significant improvement in symptom severity with a large effect size; Change in score: 13.8; 95% CI 10.6-17.0; t62 = 8.33; p < .001; d = 1.87.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• 5. Rationale for imaginal exposure and '… revisiting and recounting the traumatic memory …' (p2652), practice of imaginal exposure (15 to 45 minutes) and processing of this experience. This module can be repeated at up to 5 sessions. • 6. Imaginal exposure focusing on the '… worst moments of the trauma' (p2652). This module can be repeated for between 4 and 7 sessions. • 7. Generalising the skills developed in previous sessions and preventing relapse. • 8. Creating a 'final project' which records the trauma and documents the progress made during treatment. • Location/place of delivery – Rape crisis centre. **Comparison intervention:** Supportive counselling, a client-centred therapy which is based on the Traumagenic Dynamics Model (Finkelhor D and Browne A, 1985) and the Rogerian psychotherapy model (Rogers 1951). • Delivered by – Masters level counsellors working at a rape crisis centre who provide supportive counselling as part of their usual work.	Between group differences: The intervention group showed significantly greater improvement in symptom severity, with a large effect size: Difference in changes in score: 6.0; 95% CI 1.6-10.4; t57 = 2.67; p < .02; d = 0.81. Rates of loss of post-traumatic stress disorder diagnosis (post-traumatic stress disorder module of the DSM-IV Schedule for Affective Disorders and Schizophrenia for School-Age Children) – Intervention: Significant decrease in rates of diagnosis; p≤.001. (No further statistical data presented). Comparison: Significant decrease in rates of diagnosis; p≤.001. (No further statistical data presented). Between group differences: The intervention group showed a significantly greater decrease in rate of diagnosis, effect size not reported; Difference in changes in score: 31.1; 95% CI 14.7-34.8; t59=2.95; p=.01. Self-reported post-traumatic stress disorder symptom severity (Child PTSD Symptom Scale–Interview) - Intervention: Significant improvement in symptom severity; p≤.001. (No further statistical data presented). Comparison: Significant improvement in symptom severity; p≤.001. (No further statistical data presented). Between group differences: The intervention group showed a significantly greater improvement in symptom severity, effect size not reported; Difference in changes in score: 9.3; 95% CI 1.2-16.5; t59=2.55; p=.02. Self-reported depression severity (Children's Depression Inventory) –	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Delivered to – Female adolescents between the ages of 13 and 18 with recent experience of sexual abuse. • Duration, frequency, intensity, etc. – Up to 14 weekly sessions between 60 and 90 minutes in duration. Prior to randomisation both groups participated in preparatory sessions (up to three) which focused on case management concerns such as level of parental involvement, desire for treatment, legal processes and safety concerns (participants assessed as having active suicidal plans were excluded from the study). • Key components and objectives of intervention – The authors describe supportive counselling as a therapy which aims to establish a therapeutic relationship which is empowering, trusting and validating. Participants decide '… when, how, and whether or not to address their trauma' (p 2652). Counsellors actively listen, are empathic, encourage participants to talk about their feelings and '… express belief in the participant's ability to cope.' (p2652). Sessions are directed by participants with the exception of 4 and 8 during which counsellors ask participants to talk about their feelings regarding their trauma. This discussion and the time devoted to	Intervention: Significant improvement in symptom severity; p≤.001. (No further statistical data presented). Comparison: Significant improvement in symptom severity; p≤.001. (No further statistical data presented). Between group differences: The intervention group showed a significantly greater improvement in symptom severity, effect size not reported; Difference in changes in score: 7.2; 95% CI 1.4-13.0; t139=2.43; p=.02. Interviewer-rated functioning (Children's Global Assessment Scale) - Intervention: Significant improvement in symptom severity; p≤.001. (No further statistical data presented). Comparison: Significant improvement in symptom severity; p≤.001. (No further statistical data presented). Between group differences: The intervention group showed a significantly greater improvement in functioning, effect size not reported; Difference in changes in score: 11.2; 95% CI 4.5-17.9; t60=3.25; p=.01. Participants achieving a 'good' response to treatment (defined as a score of ≤ 8 on the Child PTSD Symptom Scale–Interview, measuring symptom severity, equivalent to ≥ 2 SDs below baseline mean) – Post-treatment Between group differences: Significantly more participants in the intervention group were classed as good responders than those in the comparison group; n=22.7 of 31, 73.3% vs. n=8.2 of 30, 27.3%; p=.001. 12 month follow-up Between group differences: Significantly more participants in the intervention group were classed as good	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	it are recorded by counsellors. The authors note that no participants described their trauma during these sessions. • Content/session titles – The first session focuses on orientation to supportive counselling but no further details on content are provided. • Location/place of delivery – Rape crisis centre. **Outcomes measured:** • Incidence of abuse and neglect - Not measured. • Risk of abuse and neglect - Not measured. • Quality of parenting and parent-child relationship – Not measured. • Children and young people's health and wellbeing – Post-traumatic stress disorder symptom severity (interviewer-rated) was measured using the Child PTSD Symptom Scale–Interview (ages 8 to 18; Foa et al. 2001). Scores range between 0 and 51 with higher scores indicating increased severity. A score of 0-10 is considered below the threshold for diagnosis; 11–15 is considered subclinical; 1–-20 = mild; 21–25 = moderate; 26–30 = moderately severe; 31–40 = severe; 41–51 = extremely severe).	responders than those in the comparison group; n=22.2 of 31, 71.7% vs. n=11.9 of 30, 39.7%; p=.02. Participants with a good response at post-treatment who maintained this at 12 month follow-up Between group differences: No significant difference (intervention=81.5% vs. comparison=70.0%; p=.53. **Narrative findings – effectiveness** **Baseline to post-treatment –** Interviewer-rated post-traumatic stress disorder symptom severity (Child PTSD Symptom Scale–Interview) - Both groups showed significant improvements in symptom severity between baseline and post-treatment, with large effect sizes; however the intervention group showed a significantly greater improvement than the comparison group, with a large effect size. Rates of loss of post-traumatic stress disorder diagnosis (post-traumatic stress disorder module of the DSM-IV Schedule for Affective Disorders and Schizophrenia for School-Age Children) – Both groups showed significant decreases in rates of diagnosis between baseline and post-treatment, however the intervention group showed a significantly greater decrease than the comparison group (effect sizes are not reported). Self-reported post-traumatic stress disorder symptom severity (Child PTSD Symptom Scale–Interview) -	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Post-traumatic stress disorder symptom severity was also measured by child self-report using the Child PTSD Symptom Scale-Self Report (ages 8 to 18, Foa et al. 2001) and uses the same scoring system as the Child PTSD Symptom Scale–Interview. Presence or absence of post-traumatic stress disorder diagnosis was measured using the post-traumatic stress disorder module of the DSM-IV Schedule for Affective Disorders and Schizophrenia for School-Age Children (Kaufman et al. 1996). Self-reported depression severity was measured using the Children's Depression Inventory (ages 7 to 17, Kovacs 1985). Scores range between 0 and 54 with higher scores indicating increased severity. A score of 0–13 is considered below the threshold for diagnosis; scores between 14 and 19 indicate the possibility of a depressive disorder; and scores between 20 and 54 indicate the presence of a depressive disorder. Functioning was measured using the Children's Global Assessment Scale (ages 4 to 18, Shaffer et al. 1983). Scores range from between 1 and 100 with lower scores indicating lower levels of functioning. Scores between 1 and 10 indicate	Both groups showed significant improvements in symptom severity between baseline and post-treatment; however the intervention group showed a significantly greater improvement than the comparison group (effect sizes are not reported). Self-reported depression severity (Children's Depression Inventory) – Both groups showed significant improvements in symptom severity between baseline and post-treatment; however the intervention group showed a significantly greater improvement than the comparison group (effect sizes are not reported). Interviewer-rated functioning (Children's Global Assessment Scale) - Both groups showed significant improvements in functioning between baseline and post-treatment; however the intervention group showed a significantly greater improvement than the comparison group (effect sizes are not reported). **Post-treatment to 12 month follow-up –** Interviewer-rated post-traumatic stress disorder symptom severity (Child PTSD Symptom Scale–Interview) – There were no significant improvements in symptom severity in either group between post-treatment and 12 month follow-up and between group differences in rate of improvement were also non-significant (effect sizes are not reported). Rates of loss of post-traumatic stress disorder diagnosis (post-traumatic stress disorder module of the	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	that the child needs supervision; a score between 51 and 60 indicates '… variable functioning with sporadic difficulties …' (p2652); and a score of 91–100 indicates '… superior functioning in all areas' (p2652). Participants achieving a 'good' response to treatment (defined as a score of ≤ 8 on the Child PTSD Symptom Scale–Interview, measuring symptom severity, equivalent to ≥ 2 SDs below baseline mean) • Caregiver/parent health and wellbeing – Not measured. • Satisfaction with services – Not measured. • Service outcomes – Not measured. **Follow-up:** Outcome measures were assessed at mid-treatment, post-treatment and at 3, 6 and 12 months post-treatment, however only data collected at baseline, post-treatment and 12 months post-treatment are reported in full. **Costs?** No. No costs or resource use information is provided.	DSM-IV Schedule for Affective Disorders and Schizophrenia for School-Age Children) – There were no significant decreases in rates of diagnosis in either group between post-treatment and 12 month follow-up and between group differences in rate of diagnosis were also non-significant (effect sizes are not reported). Self-reported post-traumatic stress disorder symptom severity (Child PTSD Symptom Scale–Interview) - There were no significant improvements in symptom severity in either group between post-treatment and 12 month follow-up and between group differences in rates of improvement were also non-significant (effect sizes are not reported). Self-reported depression severity (Children's Depression Inventory) – There were no significant improvements in symptom severity in either group between post-treatment and 12 month follow-up and between group differences in rates of improvement were also non-significant (effect sizes are not reported). Interviewer-rated functioning (Children's Global Assessment Scale) - There were no significant improvements in symptom severity in either group between post-treatment and 12 month follow-up and between group differences in rates of improvement were also non-significant (effect sizes are not reported). **Baseline to 12 month follow-up -**	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Interviewer-rated post-traumatic stress disorder symptom severity (Child PTSD Symptom Scale–Interview) – Both groups showed significant improvements in symptom severity between baseline and 12 month follow-up, with large effect sizes; however the intervention group showed a significantly greater improvement than the comparison group, with a large effect size. Rates of loss of post-traumatic stress disorder diagnosis (post-traumatic stress disorder module of the DSM-IV Schedule for Affective Disorders and Schizophrenia for School-Age Children) – Both groups showed significant decreases in rates of diagnosis between baseline and 12 month follow-up, however the intervention group showed a significantly greater decrease than the comparison group (effect sizes are not reported). Self-reported post-traumatic stress disorder symptom severity (Child PTSD Symptom Scale–Interview) – Both groups showed significant improvements in symptom severity between baseline and 12 month follow-up; however the intervention group showed a significantly greater improvement than the comparison group (effect sizes are not reported). Self-reported depression severity (Children's Depression Inventory) – Both groups showed significant improvements in symptom severity between baseline and 12 month follow-up; however the intervention group showed a significantly greater improvement than the comparison group (effect sizes are not reported).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Interviewer-rated functioning (Children's Global Assessment Scale) - Both groups showed significant improvements in functioning between baseline and 12 month follow-up; however the intervention group showed a significantly greater improvement than the comparison group (effect sizes are not reported). Participants achieving a 'good' response to treatment (defined as a score of ≤8 on the Child PTSD Symptom Scale–Interview, measuring symptom severity, equivalent to ≥2 SDs below baseline mean) – Post-treatment At post-treatment, significantly more participants in the intervention group were classed as good responders than those in the comparison group. 12 month follow-up - At 12 month follow-up, significantly more participants in the intervention group were classed as good responders than those in the comparison group. Participants with a good response at post-treatment who maintained this at 12 month follow-up The proportion of participants who were classed as good responders and maintained this score at 12 month follow-up did not differ significantly by group.	

5. Goldman Fraser J, Lloyd SW, Murphy RA et al. (2013) Child exposure to trauma: Comparative effectiveness of interventions addressing maltreatment. Comparative Effectiveness Review 89: 1–161

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: 1. To evaluate the comparative	**Participants:** The population for the review as a whole were a) Children and young people:	**Statistical data for studies relevant to question 16:**	**Overall assessment of internal validity:** ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
efficacy and effectiveness of psychosocial and pharmacological interventions that address child wellbeing and/or promote positive child welfare outcomes (safety, placement stability, and permanency) for maltreated children ages birth to 14 years. 2. To assess the comparative effectiveness of interventions (a) with different treatment characteristics, (b) for child and caregiver subgroups, and (c) for engaging and retaining children and/or caregivers in treatment; and (3) To assess harms associated with interventions for this population. **Methodology:** Used review methods described in AHRQ's Methods Guide for Effectiveness and Comparative Effectiveness Reviews.	1. Aged 0–14 years of age, exposed to maltreatment (defined as child abuse [acts of commission: words or overt actions that cause harm, potential harm, or threat of harm to a child] and child neglect [acts of omission: failure to provide for a child's basic physical, emotional, or educational needs or to protect a child from harm or potential harm]). 2. Children of the same ages involved with the child welfare system (including foster care), and caregivers of maltreated children when they were the target of an intervention. 3. Children with known CPS involvement b) Caregivers and families Also included primary caregiver(s) caregivers of maltreated children when they were the target of an intervention. **Sample characteristics:** In the studies for which we have conducted data extraction for this question, the sample characteristics were as follows: • Age - Jaberghaderi et al. 2004 – not included in review. Trowell et al. 2002 – 6 to 14 years McGain and McKinzey 1995 – not included Cohen et al. 1996 – 2.11 to 7.1 years	This systematic review examined 4 interventions aimed at young people who had experienced sexual abuse: eye movement desensitisation and reprocessing, group psychotherapy, group treatment programme, trauma-focused CBT. 1. Eye movement desensitisation and reprocessing One study included (Jaberghaderi et al. 2004), but this was conducted in a country which is out of scope of the current review (Iran), and therefore the results are not recorded here. 2. Group psychotherapy One UK RCT identified (Trowell et al. 2002) evaluating a medium-intensity, psychoeducational and psychotherapeutic group treatment for sexually abused girls compared with an active control (conventional psychoanalytic individual therapy). 2.1 Incidence of abuse and neglect - Not measured. 2.2 Risk of abuse and neglect - Not measured. 2.3 Quality of parenting and parent-child relationships - Not measured. 2.4 Children and young people's health and wellbeing Re-experiencing traumatic events scale (Orvaschel PTSD scale): T1: Poorer outcome for intervention compared to active control (d=0.60, p=not reported, significant) T2: Poorer outcome for intervention compared to active control (d=0.79, p=not reported, significant)	**Overall assessment of external validity:** + **Overall validity rating:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Data analysed in the form of a qualitative synthesis. A quantitative meta-analysis was not performed due to issues of heterogeneity, insufficient numbers of similar studies, and poor outcome reporting. **Country:** The studies relevant to Q16 were conducted in the UK (1 study). **Source of funding:** Government. The Agency for Healthcare Research and Quality (AHRQ), Evidence-based Practice Centers (EPCs), USA.	Cohen et al. 2004 – 8 to 14.11 years Deblinger et al. 2001 – 2 to 8 years • Sex - Jaberghaderi et al. 2004 – not included in review. Trowell et al. 2002 – all female McGain and McKinzey 1995 – not included in review Cohen et al. 1996 – not reported Cohen et al. 2004 – not reported Deblinger et al. 2001 – not reported • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse – The review included studies/interventions for sexually abused children. • Looked after or adopted status – The review included children who remained in the care of their biological parent as well as those in out-of-home care (e.g., foster care, kinship care, group home care). • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:**	Persistent avoidance of stimuli: (Orvaschel PTSD scale): T1: Poorer outcome for intervention compared to active control (d=0.66, p=not reported, significant) T2: Poorer outcome for intervention compared to active control (d=0.36, p=not reported, significant) Persistent symptoms of increased arousal: (Orvaschel PTSD scale): T1: No difference between intervention and active control (p=not reported, non-significant) T2: No difference between intervention and active control (p=not reported, non-significant) Impairment index (Kiddie Global Assessment Scale): Q1: No difference between intervention and active control (p=not reported, non-significant) Q2: No difference between intervention and active control (p=not reported, non-significant) 2.5 Caregiver health and wellbeing - Not measured. 2.6 Satisfaction with services - Not measured. 2.7 Service outcomes - Not measured. 3. Group programme for sexual abuse (McGain and McKinzey 1995). NB It is not clear how this intervention differs from the group psychotherapy intervention above). The review found one non-randomised control trial. As this is not an RCT, data has not been extracted for this study. 4. Trauma-focused CBT	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Systematic reviews - number of studies Qualitative synthesis of 11 trials (9 RCTs and 1 cohort study) on physical abuse and neglect. Studies with a high risk of bias are excluded by the authors in the results. **Intervention:** 1. Eye movement desensitisation and reprocessing – not included in review 2. Group psychotherapy - A psychoeducational and psychotherapeutic group treatment programme of medium-intensity (up to 18, 50 minute sessions with concurrent parent sessions every 2 weeks – group-based). The control treatment is described as high intensity (up to 30, 50 minute sessions every week). Both treatments included '… generic and abuse-specific components …' (p11) such as establishing a therapeutic relationship and management of anxiety, and a '… caregiver-directed component comprising social work support (delivered in either a group or individual mode aligned with that of the child-directed component)' (p57). 3. Group programme for sexual abuse – not included in review 4. Trauma-focused CBT - therapy which aims to reduce maladaptive responses to sexual abuse exposure or	The review identified three RCTs evaluating the efficacy of Trauma-Focused Cognitive Behavioural Therapy (TF-CBT) (Cohen et al. 1996; Cohen et al. 2004; Deblinger et al. 2001). 4.1 Incidence of abuse and neglect - Not measured in any studies. 4.2 Risk of abuse and neglect - Not measured in any studies. 4.3 Quality of parenting and parent-child relationships Cohen et al. 1996 - Not measured Cohen et al. 2004 - Improved parenting practices (Parenting Practices Questionnaire): Intervention group significantly better than control (d=0.57, p<0.001) Deblinger et al. 2001 - Improved parenting practices (Parenting Practices Questionnaire): No difference between intervention and control (p=not reported, non-significant) 2.4 Children and young people's health and wellbeing - Cohen et al. 1996 'Improvements in social competence (Child Behaviour Checklist Social Competence): No difference between intervention and control (p=not reported, non-significant). Improvements in behaviour (Child Behaviour Checklist Behavioural Profile - Total): Intervention group significantly better than control (p<0.01).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	other traumatic events. The key objectives of treatment are to enhance the child's ability to express feeling; recognise the relationship between behaviours, feelings and thoughts; and to develop coping skills. The programmes involve 'gradual exposure' or creation of the trauma narrative, and cognitive processing of traumatic event, psychoeducation in relation to child sexual abuse and body safety, and support for parents in relation to behavioural management. Participants received between 12 and 16 weekly sessions which lasted for between one and 1 and a half hours. Sessions were delivered individually to both children and parents, and jointly to children and parents. **Outcomes measured:** • Jaberghaderi et al. 2004 – not included in review. • Trowell et al. 2002 – 2.4 Children and young people's health and wellbeing measured using re-experiencing traumatic events scale (Orvaschel PTSD scale); persistent avoidance of stimuli: (Orvaschel PTSD scale); persistent symptoms of increased arousal: (Orvaschel PTSD scale); Impairment index (Kiddie Global Assessment Scale). • McGain and McKinzey 1995 – not included in review	Improvements in internalising symptoms (Child Behaviour Checklist - Internalising): Intervention group significantly better than control (p>0.002). Improvements in externalising symptoms (Child Behaviour Checklist - Externalising): No difference between intervention and control (p=not reported, non-significant). Improvements in sexual behaviours (Child Sexual Behaviour Inventory): Intervention group significantly better than control (p>0.05)' Cohen et al. 2004 'Decrease in re-experiencing of traumatic event (Schedule for Affective Disorders and Schizophrenia for School-Aged Children Present and Lifetime Version - Re-experiencing): Intervention group significantly better than control (d=0.49, p<0.01). Decrease in avoidance of reminders of traumatic event (Schedule for Affective Disorders and Schizophrenia for School-Aged Children Present and Lifetime Version - Avoidance): Intervention group significantly better than control (d=0.70, p<0.0001). Decrease in hypervigilance ((Schedule for Affective Disorders and Schizophrenia for School-Aged Children Present and Lifetime Version - Hypervigilance): Intervention group significantly better than control (d=0.40, p<0.01). Improvements in behaviour (Child Behavior Checklist Total) Intervention group significantly better than control (d=33, p<0.01)	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Cohen et al. 1996 – Child health and wellbeing - improvements in social competence (Child Behaviour Checklist Social Competence); internalising and externalising symptoms (Child Behaviour Checklist), sexualised behaviours (Child Sexual Behaviour Inventory). • Cohen et al. 2004 – Quality of parenting and parent-child relationships: Improved parenting practices (Parenting Practices Questionnaire) Child health and wellbeing –re-experiencing, avoidance, hypervigilance (Schedule for Affective Disorders and Schizophrenia for School-Aged Children Present and Lifetime Version); behaviour, social competence (Child Behaviour Checklist); depression, sexual behaviours (Children's Depression Inventory); anxiety and proneness to anxiety (State-Trait Anxiety Inventory for Children) Caregiver health and wellbeing – parent self-report of depression (Beck Depression Inventory) • Deblinger et al. 2001 – Quality of parenting and parent-child relationships: Improved parenting practices (Parenting Practices Questionnaire).	Improvements in social competence (Child Behavior Checklist Competence) No difference between intervention and control (p=not reported, non-significant) Improvements in internalising problems (Child Behaviour Checklist Internalising): No difference between intervention and control (p=non-significant) Improvements in externalising behaviour (Child Behaviour Checklist Externalising): No difference between intervention and control (p=non-significant). Improvements in depression (Children's Depression Inventory) Intervention group significantly better than control (d=0.30, p<0.05) Improvements in sexual behaviours (Children's Depression Inventory [sic?]): No difference between intervention and control (p=not reported, non-significant) Improvements in proneness to anxiety (State-Trait Anxiety Inventory for Children Trait): No difference between intervention and control (p=not reported, non-significant) Improvements in fleeting anxiety (State-Trait Anxiety Inventory for Children State): No difference between intervention and control (p=not reported, non-significant)' Deblinger et al. 2001 'Changes in PTSD symptoms: No difference between intervention and control (p=not reported, non-significant)	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Child health and wellbeing – PTSD symptoms, behaviour (Child Behaviour Checklist), sexual behaviour (Child Sexual Behaviour Inventory) Caregiver health and wellbeing – Maternal PTSD symptoms (Symptom Checklist-90-Revised), maternal distress (Impact of events scale).	Changes in behaviour (Child Behaviour Checklist): No difference between intervention and control (p=not reported, non-significant) Changes in sexual behaviours (Child Sexual Behaviour Inventory) No difference between intervention and control (p=not reported, non-significant)' 4.5 Caregiver health and wellbeing: Cohen et al. 1996 - Not measured Cohen et al. 2004 - Parent self-report of depression (Beck Depression Inventory): Intervention group significantly better than control (d=0.38, p<0.05) Deblinger et al. 2001 - 'Maternal PTSD symptoms (Symptom Checklist-90-Revised): No difference between intervention and control (p=not reported, non-significant) Maternal distress-intrusive thoughts (Impact of events scale): Intervention group significantly better than control (p<0.05, d=not reported) Maternal distress - avoidant thoughts (Impact of events scale): No difference between intervention and control (p=not reported, non-significant) Satisfaction with services: Not measured in any studies. Service outcomes: Not measured in any studies. **Narrative findings for studies relevant to Q16:** 1. Eye movement desensitisation and reprocessing	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		One study included (Jaberghaderi et al. 2004), but this was conducted in a country which is out of scope of the current review (Iran), and therefore the results are not recorded here. 2. Group psychotherapy One UK RCT identified (Trowell et al. 2002) evaluating a medium-intensity, psychoeducational and psychotherapeutic group treatment for sexually abused girls compared with an active control (conventional psychoanalytic individual therapy). The study did not measure incidence of abuse and neglect; risk of abuse and neglect; the quality of parenting and parent-child relationships; caregiver health and wellbeing; satisfaction with services; or service outcomes. The study did measure children and young people's health and wellbeing and found that: Post-traumatic stress disorder – re-experiencing - At 12 month follow-up participants randomised to the group psychotherapy plus caregiver support group had worse outcomes in comparison to participants randomised to the individual therapy plus caregiver support group with a medium effect size d=0.60 (significance value not reported, measured using Overaschel's Re-experience of the traumatic event subscale of the Post-traumatic stress disorder scale of the Schedule for affective disorders and schizophrenia for school—age children present and lifetime version). Participants in this group also had worse outcomes on this measure at 24 month follow-up in comparison to the control group with a medium to large effect size d=0.79 (significance value not reported). Post-traumatic stress disorder – re-experiencing -	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		At 12 month follow-up participants randomised to the individual psychotherapy plus caregiver support group had better outcomes in comparison to participants randomised to the group therapy plus caregiver support group with a medium effect size d=0.60 (significance value not reported but reported as significant by the review authors, measured using Overaschel's Re-experience of the traumatic event subscale of the Post-traumatic stress disorder scale of the Schedule for affective disorders and schizophrenia for school—age children present and lifetime version). This difference remained significant at 24 months follow-up with a medium to large effect size (d=0.79; significance value not provided but reported as significant by the review authors). Post-traumatic stress disorder – persistent avoidance - At 12 month follow-up participants randomised to the individual psychotherapy plus caregiver support group had better outcomes in comparison to participants randomised to the group therapy plus caregiver support group with a medium to large effect size (d=0.66, significance value not reported, measured using Overaschel's Persistent avoidance of stimuli subscale of the Post-traumatic stress disorder scale of the Schedule for affective disorders and schizophrenia for school—age children present and lifetime version). This difference remained significant at 24 months follow-up with a small to medium effect size (d=0.36; significance value not provided but reported as significant by the review authors).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Persistent symptoms of increased arousal – At 12 month and 24 month follow-ups there were no significant differences between individual and group psychotherapy conditions in persistent symptoms of increased arousal measured using the Orvaschel PTSD scale (effect sizes and significance values not provided, reported as non-significant by review authors). Impairment - At 12 month and 24 month follow-ups there were no significant differences between individual and group psychotherapy conditions in scores of impairment measured using the Kiddie Global Assessment Scale (effect sizes and significance values not provided, reported as non-significant by review authors) 3. Group program for sexual abuse (McGain and McKinzey 1995) NB It is not clear how this intervention differs from the group psychotherapy intervention above. The review found one non-randomised control trial. As this is not an RCT, data has not been extracted for this study. 4. Trauma-focused CBT The review identified three RCTs evaluating the efficacy of Trauma-Focused Cognitive Behavioural Therapy (TF-CBT) (Cohen et al. 1996; Cohen et al. 2004; Deblinger et al. 2001). None of these studies measured incidence of abuse and neglect; risk of abuse and neglect; satisfaction with services; or service outcomes. Quality of parenting and parent-child relationships was measured by Cohen et al. 2004 and Deblinger et al. 2001.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Cohen et al. 2004 found that there was a significant difference between groups with a medium effect size in relation to improved parenting practices measured using the (Parenting Practices Questionnaire), with participants randomised to the intervention group achieving higher scores than those randomised to the control group. Deblinger et al. 2001 found that there was no significant differences between groups in scores on this measure (p values not provided, reported as non-significant by review authors). Children and young people's health and wellbeing was measured by Cohen et al. 1996; Cohen et al. 2004; and Deblinger et al. 2001. Cohen et al. 1996 found that there were no significant difference between the intervention and control groups in improvements in social competence (measured using the Child Behaviour Checklist Social Competence); and improvements in externalising symptoms (measured using the Child Behaviour Checklist - Externalising). NB effect sizes and p values not provided, reported as non-significant by review authors. The intervention group showed significantly greater improvements in behaviour measured using total scores on the Child Behaviour Checklist Behavioural Profile ($p<0.01$); significantly greater improvements in internalising symptoms (measured using the Child Behaviour Checklist – Internalising; $p>0.002$; and significantly greater improvements in sexual behaviours (measured using the Child Sexual Behaviour Inventory ($p>0.05$). NB Effect sizes not provided.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Cohen et al. 2004 found that found that there were no significant difference between the intervention and control groups in improvements in social competence (measured using the Child Behaviour Checklist – Competence; improvements in internalising problems (measured using the Child Behaviour Checklist - Internalising); improvements in sexual behaviours (measured using the Children's Depression Inventory [sic]); improvements in proneness to anxiety (measured using the State-Trait Anxiety Inventory for Children – Trait); improvements in fleeting anxiety (measured using the State-Trait Anxiety Inventory for Children – State); and improvements in externalising behaviour (measured using the Child Behaviour Checklist - Externalising. NB Effect sizes and significance values not provided, reported as non-significant by review authors. In contrast, the intervention group showed significantly greater improvements in behaviour measured using total scores on the Child Behavior Checklist – Total; d=0.33, p<0.01); and significantly greater improvements in depression (measured using the Children's Depression Inventory; d=0.30, p<0.05.) The intervention group also showed significantly greater decreases in re-experiencing of traumatic events (measured using the Schedule for Affective Disorders and Schizophrenia for School-Aged Children Present and Lifetime Version - Re-experiencing; d=0.49, p<0.01; significantly greater decreases in avoidance of reminders of traumatic event (measured using the Schedule for Affective Disorders and Schizophrenia for School-Aged Children Present and Lifetime Version – Avoidance; d=0.70, p<0.0001); and significantly greater decreases in hypervigilance (measured	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		using the Schedule for Affective Disorders and Schizophrenia for School-Aged Children Present and Lifetime Version – Hypervigilance; d=0.40, p<0.01) Deblinger et al. 2001 found that there were no significant differences between groups in relation to changes in post-traumatic stress disorder symptoms; changes in behaviour (measured using the Child Behaviour Checklist; or changes in sexual behaviours (measured using the Child Sexual Behaviour Inventory). NB Effect sizes and significance values not provided, reported as non-significant by review authors. Caregiver health and wellbeing was measured by Cohen et al. 2004 and Deblinger et al. 2001. Cohen et al. 2004 found that parents whose children were randomised to the intervention group had significantly better levels of self-reported depression (measured using the Beck Depression Inventory; d=0.38, p<0.05). Deblinger et al. 2001 found that mothers whose children were randomised to the intervention group had significantly better levels of maternal distress/intrusive thoughts (measured using the Impact of events scale (p<0.05, d=not reported) than participants whose children were randomised to the control group. However there were no significant differences found between groups on measures of maternal distress/avoidant thoughts (measured using the Impact of events scale) and maternal post-traumatic stress disorder symptoms (measured using the Symptom Checklist-90-Revised). NB Effects sizes and significance values not provided, reported as significant by review authors.	

6. Leenarts, LE, Diehle J, Doreleijers TA et al. (2013) Evidence-based treatments for children with trauma-related psychopathology as a result of childhood maltreatment: A systematic review. European Child and Adolescent Psychiatry 22: 269–83

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To '… systematically evaluate psychotherapeutic treatments for children exposed to childhood maltreatment and to describe treatments which focus on the above mentioned broad range of psychopathological outcomes' (p270). **Methodology:** Systematic review of randomised and non-randomised controlled trials. There are eight studies which are relevant to question 16 (randomised controlled trials evaluating interventions designed to respond to sexual abuse). **Country:** The included studies were conducted in a range of countries; and although one was conducted in the UK (Trowell et al. 2002)	**Participants:** • Children and young people. The review inclusion criteria stipulated that studies had to include children between the ages of 6 and 18 years who had experienced maltreatment. Studies involving children who had experienced '… war related violence or traumatic grief …' were excluded. The NCCSC has only extracted data in relation to studies which have a sample composed entirely of sexually abused children (none of the included studies with a mixed sample report subgroup analyses). • Caregivers and families - Inclusion criteria specified that studies evaluating interventions delivered only to a maltreating parent would only be included if post-traumatic stress symptoms of the child were measured. Although it is not reported specifically by the review authors a number of the included studies appear to have also involved a parent or caregiver. Those which are relevant to question 16 (responses to sexual abuse) are: Cohen et al. 2004; Cohen et al. 2005; Deblinger et al. 2001; Deblinger et al. 2011; and King et al. 2000.	**Statistical data – Children and young people's health and wellbeing outcomes:** NB. It is not clear whether the review authors calculated effect sizes or whether those provided are quoted from the included studies. The authors do not specify that effect sizes are Cohen's d however, their description suggests that this is the case. Significance levels and direction of effects are not reported although the narrative findings sometimes refer to statistical significance; however due to concerns regarding the accuracy of this, this information has not been extracted. The review included 9 studies relevant to question 16 and the reported effect sizes ranged from very small to large (0.07 to 1.24). **Cohen et al., 2004 (N=229 randomised) –** Post-traumatic stress disorder (measured using the Post-traumatic stress disorder supplement [re-experiencing] of the Schedule for affective disorders and schizophrenia for school-age children present and lifetime version) – Post-test: There was a small to medium difference between participants randomised to the Trauma-Focused Cognitive Behavioural Therapy group and those randomised to the Child-Centred Therapy group (significance value not reported); 0.49. Post-traumatic stress disorder (measured using the Post-traumatic stress disorder supplement [avoidance] of the Schedule for affective disorders and schizophrenia for school—age children present and lifetime version) –	**Overall assessment of internal validity:** - **Overall assessment of external validity:** ++ **Overall assessment of validity:** -

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
the majority were from the USA. Only data relating to studies which meet the geographical criteria specified in the NCCSC protocol and are relevant to question 16 have been extracted. Cohen et al. 2004 – USA. Cohen et al. 2005 – USA. Danielson et al. 2012 – USA. Deblinger et al. 2001 – USA. Deblinger at al. 2011 – USA. King et al. 2000 – USA. Trowell et al. 2002 – UK. **Source of funding:** Other – unclear - LSG-Rentray, Residential and Ambulant Treatment Center for Children and Adolescents, the Netherlands.	**Sample characteristics:** • Age - The review protocol states that studies focusing on children between the ages of 6 and 18 were to be included, however the ages of participants in the included studies ranged from two to 25. No details on the ages of participating parents are included. The age range of samples in those studies relevant to the question 16 are – Cohen et al., 2004 – 8–14 years. Cohen et al., 2005 – 7–14 years. Danielson et al., 2012 – 13–17 years. Deblinger et al., 2001 – 2–8 years. Deblinger at al., 2011 – 4–11 years. King et al., 2000 – 5–17 years. Trowell et al., 2002 – 6–14 years. • Sex - The review does not provide details on the gender of participating parents despite the fact that they note the majority of studies '… involved parents or caregivers in some way' (p280). The reported gender balance of those studies relevant to the question 16 are - Cohen et al. 2004 – 21% male (attended three or more treatment sessions). Cohen et al. 2005 – 32% male. Danielson et al. 2012 – 12% male. Deblinger et al. 2001 – 39% male. Deblinger at al. 2011 – 39% male (completers).	Post-test: There was a medium to large difference between participants randomised to the Trauma-Focused Cognitive Behavioural Therapy group and those randomised to the Child-Centred Therapy group (significance value not reported); 0.70. Post-traumatic stress disorder (measured using the Post-traumatic stress disorder supplement [hyper-vigilance] of the Schedule for affective disorders and schizophrenia for school—age children present and lifetime version) – Post-test: There was a small to medium difference between participants randomised to the Trauma-Focused Cognitive Behavioural Therapy group and those randomised to the Child-Centred Therapy group (significance value not reported); 0.40. **Cohen et al., 2005 (n=82 randomised) -** Post-traumatic stress disorder (measured using the post-traumatic stress disorder subscale of the Trauma symptom checklist for children) – Post-test: There was a small difference between participants randomised to the Trauma-Focused Cognitive Behavioural Therapy group and those randomised to the Non-directive supportive Therapy group (significance value not reported); 0.22. **Danielson et al. 2012 (n=30 randomised) –** Post-traumatic stress disorder (measured using the University of California at Los Angeles post-traumatic stress disorder index for adolescents) – Post-test: There was a small to medium difference between participants randomised to the Risk Reduction through Family Therapy group and those randomised to the	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	King et al., 2000 – 31% male. Trowell et al., 2002 – 0% male. • Ethnicity - Not reported in detail for any of the included studies relevant to the NCCSC review, however, the authors note that the sample in Cohen et al., 2004 was not ethnically diverse. • Religion/belief - Not reported for any of the included studies. • Disability - Not reported for any of the included studies. • Long term health condition - Not reported for any of the included studies. • Socioeconomic position - Not reported for any of the included studies. • Type of abuse – The samples of all included studies relevant to question 16 are described as 'sexually abused children' (no further details provided) with the exceptions of Danielson et al. 2012, which the review authors describe as 'sexually assaulted children'. • Looked after or adopted status - Not reported for any of the included studies. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported for any of the included studies.	treatment as usual group (significance value not reported); 0.38. **Deblinger et al. 2001 (n=67 randomised) –** Post-traumatic stress disorder (scale used unclear) – Post-test: There was a very small difference between participants randomised to the cognitive-behavioural therapy group and those randomised to the supportive counselling group (significance value not reported); 0.07. **Deblinger et al. 2004 (n=210 randomised, experimental and comparison conditions are not clear) –** Post-traumatic stress disorder (measured using the Post-traumatic stress disorder supplement [re-experiencing] of the Schedule for affective disorders and schizophrenia for school-age children present and lifetime version) – Post-test: There was a small between group difference (mean, significance value not reported); 0.35. Post-traumatic stress disorder (measured using the Post-traumatic stress disorder supplement [avoidance] of the Schedule for affective disorders and schizophrenia for school—age children present and lifetime version) – Post-test: There was a small between group difference (mean, significance value not reported); 0.35. Post-traumatic stress disorder (measured using the Post-traumatic stress disorder supplement [hyper-vigilance] of the Schedule for affective disorders and schizophrenia for school-age children present and lifetime version) –	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Sample size: • Comparison numbers: The review authors do not report the number of participants assigned to each group for any of the included studies. • Intervention numbers: The review authors do not report the number of participants assigned to each group for any of the included studies. • Sample size (totals): Cohen et al. 2004 – n=229 randomised. Cohen et al. 2005 – n=82 randomised. Danielson et al. 2012 – n=30 randomised. Deblinger et al. 2001 – n=67 randomised. Deblinger at al. 2011 – N=210 randomised. King et al. 2000 – n=36 randomised. Trowell et al. 2002 – n=75 randomised. • Systematic reviews: The review included 27 studies reporting on a total of 26 trials. Seven of these studies are relevant to question 16 of the NCCSC review (responses to sexual abuse). **Intervention:** • Intervention category – Other – The review focuses on psychotherapeutic treatments.	Post-test: There was a small between group difference (mean, significance value not reported); 0.23. **King et al. 2000 (n=36 randomised) –** Post-traumatic stress disorder (measured using the post-traumatic stress disorder section of the Anxiety disorder interview schedule) – Post-test - child cognitive behavioural therapy vs family cognitive behavioural therapy: There was a small difference between participants randomised to the child cognitive behavioural therapy group and those randomised to the family cognitive behavioural therapy group (significance value not reported); 0.23. Post-test - child cognitive behavioural therapy vs WLC: There was a large difference between participants randomised to the child cognitive behavioural therapy group and those randomised to the waitlist control group (significance value not reported); 1.09. Post-test - family cognitive behavioural therapy vs. WLC: There was a large difference between participants randomised to the family cognitive behavioural therapy group and those randomised to the waitlist control group (significance value not reported); 1.24. **Trowell et al., 2002 (n=75 randomised) –** Post-traumatic stress disorder (measured using Overaschel's Re-experience of the traumatic event subscale of the Post-traumatic stress disorder scale of the Schedule for affective disorders and schizophrenia for school-age children present and lifetime version) - 12 month follow-up: There was a medium difference between participants randomised to the individual	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Describe intervention - The included studies relevant to question 16 of the NCCSC review evaluated: Cohen et al. 2004 – Trauma-focused cognitive behavioural therapy. Cohen et al. 2005 – Trauma-focused cognitive behavioural therapy. Danielson et al. 2012 – Risk reduction through family therapy. Deblinger et al. 2001 – Cognitive-behavioural therapy. Deblinger at al. 2011 – Trauma-focused cognitive behavioural therapy with or without trauma narrative in 8 vs. 16 sessions. King et al. 2000 – Child cognitive-behavioural therapy or family cognitive-behavioural therapy. Trowell et al. 2002 – Individual psychotherapy plus caregiver support. • Delivered by – Not reported for any of the included studies relevant to question 16 of the NCCSC review. • Delivered to - The samples of all included studies relevant to question 16 are described as 'sexually abused children' (no further details provided) with the exception of Danielson et al. 2012, the sample of which the review authors describe as 'sexually assaulted children'. Cohen et al. 2004; Cohen et al. 2005; Deblinger et al. 2001;	therapy plus caregiver support group and those randomised to the group psychotherapy plus caregiver support group (significance value not reported); 0.60. Post-traumatic stress disorder (measured using Overaschel's Persistent avoidance of stimuli subscale of the Post-traumatic stress disorder scale of the Schedule for affective disorders and schizophrenia for school-age children present and lifetime version) - 12 month follow-up: There was a medium to large difference between participants randomised to the individual therapy plus caregiver support group and those randomised to the group psychotherapy plus caregiver support group (significance value not reported); 0.66. **Narrative findings - Children and young people's health and wellbeing outcomes:** The review included nine studies relevant to question 16 and the reported effect sizes ranged from very small to large (0.07 to 1.24). **Cohen et al. 2004 (n=229 randomised) –** Post-traumatic stress disorder (measured using the Post-traumatic stress disorder supplement [re-experiencing] of the Schedule for affective disorders and schizophrenia for school-age children present and lifetime version) – Post-test: There was a small to medium difference between participants randomised to the Trauma-Focused Cognitive Behavioural Therapy group and those randomised to the Child-Centred Therapy group (significance value not reported).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Deblinger et al., 211; and King et al. 2000 all appear to have also included parents or caregivers. • Duration, frequency, intensity, etc. – Not reported for any of the included studies relevant to question 16 with the exception of Deblinger et al. 2011 which appears to have been delivered for either eight or 16 sessions. • Key components and objectives of intervention – Not reported for any of the included studies relevant to question 16. • Content/session titles - Not reported for any of the included studies relevant to question 16. • Location/place of delivery - Not reported for any of the included studies relevant to question 16. • Describe comparison intervention - Cohen et al. 2004 – Child-Centered Therapy. Cohen et al. 2005 – Non-directive supportive therapy. Danielson et al., 2012 – Treatment as usual. Deblinger et al., 2001 – Supportive counselling Deblinger at al. 2011 – Trauma-focused cognitive behavioural therapy with or without trauma narrative in eight vs 16 sessions. King et al. 2000 – Waitlist control.	Post-traumatic stress disorder (measured using the Post-traumatic stress disorder supplement [avoidance] of the Schedule for affective disorders and schizophrenia for school-age children present and lifetime version) – Post-test: There was a medium to large difference between participants randomised to the Trauma-Focused Cognitive Behavioural Therapy group and those randomised to the Child-Centred Therapy group (significance value not reported). Post-traumatic stress disorder (measured using the Post-traumatic stress disorder supplement [hyper-vigilance] of the Schedule for affective disorders and schizophrenia for school—age children present and lifetime version) – Post-test: There was a small to medium difference between participants randomised to the Trauma-Focused Cognitive Behavioural Therapy group and those randomised to the Child-Centred Therapy group (significance value not reported). **Cohen et al. 2005 (n=82 randomised) -** Post-traumatic stress disorder (measured using the post-traumatic stress disorder subscale of the Trauma symptom checklist for children) – Post-test: There was a small difference between participants randomised to the Trauma-Focused Cognitive Behavioural Therapy group and those randomised to the Non-directive supportive Therapy group (significance value not reported). **Danielson et al. 2012 (n=30 randomised) –**	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Trowell et al. 2002 – Group psychotherapy plus caregiver support. • Delivered by - Not reported for any of the included studies relevant to question 16. • Delivered to - The samples of all included studies relevant to question 16 are described as 'sexually abused children' (no further details provided) with the exception of Danielson et al. 2012, the sample of which the review authors describe as 'sexually assaulted children'. Cohen et al. 2004; Cohen et al. 2005; Deblinger et al. 2001; Deblinger et al. 2011; and King et al. 2000 all appear to have also included parents or caregivers. • Duration, frequency, intensity, etc. - Not reported for any of the included studies relevant to question 16. • Key components and objectives of intervention - Not reported for any of the included studies relevant to question 16. • Content/session titles - Not reported for any of the included studies relevant to question 16. • Location/place of delivery - Not reported for any of the included studies relevant to question 16. **Outcomes measured:** On the whole the review is unclear about the outcomes measured by the individual	Post-traumatic stress disorder (measured using the University of California at Los Angeles post-traumatic stress disorder index for adolescents) – Post-test: There was a small to medium difference between participants randomised to the Risk Reduction through Family Therapy group and those randomised to the treatment as usual group (significance value not reported). **Deblinger et al. 2001 (n=67 randomised) –** Post-traumatic stress disorder (scale used unclear) – Post-test: There was a very small difference between participants randomised to the cognitive-behavioural therapy group and those randomised to the supportive counselling group (significance value not reported). **Deblinger et al. 2004 (n=210 randomised, experimental and comparison conditions are not clear) –** Post-traumatic stress disorder (measured using the Post-traumatic stress disorder supplement [re-experiencing] of the Schedule for affective disorders and schizophrenia for school-age children present and lifetime version) – Post-test: There was a small between group difference (mean, significance value not reported); 0.35. Post-traumatic stress disorder (measured using the Post-traumatic stress disorder supplement [avoidance] of the Schedule for affective disorders and schizophrenia for school-age children present and lifetime version) – Post-test: There was a small between group difference (mean, significance value not reported); 0.35.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	studies and provides little detail on how the review itself used this information (e.g. whether outcomes were included in the search strategy). The individual studies may have measured more outcomes than those listed below and effect sizes are not always included. Total scores were reported by the review authors where these were available. • Incidence of abuse and neglect - This review did not record whether the included studies measured incidence of abuse and neglect. • Risk of abuse and neglect - This review did not record whether the included studies measure risk of abuse and neglect. • Quality of parenting and parent-child relationships - This review did not record whether the included studies measured quality of parenting and parent-child relationships. • Children and young people's health and wellbeing outcomes - The authors simply note that the review aims to evaluate treatments which focus on a 'broad range of psychopathological outcomes' such as anxiety, post-traumatic stress disorder, suicidal ideation, and substance abuse. However, effect sizes are only presented for post-traumatic stress disorder related	Post-traumatic stress disorder (measured using the Post-traumatic stress disorder supplement [hyper-vigilance] of the Schedule for affective disorders and schizophrenia for school-age children present and lifetime version) – Post-test: There was a small between group difference (mean, significance value not reported); 0.23. **King et al. 2000 (n=36 randomised) –** Post-traumatic stress disorder (measured using the post-traumatic stress disorder section of the Anxiety disorder interview schedule) – Post-test - child cognitive behavioural therapy vs family cognitive behavioural therapy: There was a small difference between participants randomised to the child cognitive behavioural therapy group and those randomised to the family cognitive behavioural therapy group (significance value not reported). Post-test - child cognitive behavioural therapy vs WLC: There was a large difference between participants randomised to the child cognitive behavioural therapy group and those randomised to the waitlist control group (significance value not reported). Post-test - family cognitive behavioural therapy vs WLC: There was a large difference between participants randomised to the family cognitive behavioural therapy group and those randomised to the waitlist control group (significance value not reported). **Trowell et al. 2002 (n=75 randomised) –** Post-traumatic stress disorder (measured using Overaschel's Re-experience of the traumatic event subscale of the Post-traumatic stress disorder scale of	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	scales. Very little detail is presented on these and it is not always clear if these are self-report measures. Measures used in included studies relevant to the NCCSC review were - Post-traumatic stress disorder was measured using the re-experiencing; hyper-vigilance; and avoidance sections of the post-traumatic stress disorder supplement of the Schedule for affective disorders and schizophrenia for school-age children present and lifetime version (K-SADS-PL) - used in Cohen et al. 2004; Deblinger et al., 2011. Post-traumatic stress disorder was measured using the post-traumatic stress disorder subscale of the Trauma Symptom Checklist for Children (TSCC) - used in Cohen et al. 2005. Post-traumatic stress disorder was measured using the University of California at Los Angeles post-traumatic stress disorder index for adolescents (UCLA PTSD-A) - used in Danielson et al. 2012. Post-traumatic stress disorder was measured using the post-traumatic stress disorder section of the Anxiety disorder interview schedule (ADIS) - used in King et al. 2000.	the Schedule for affective disorders and schizophrenia for school-age children present and lifetime version) - 12 month follow-up: There was a medium difference between participants randomised to the individual therapy plus caregiver support group and those randomised to the group psychotherapy plus caregiver support group (significance value not reported). Post-traumatic stress disorder (measured using Overaschel's Persistent avoidance of stimuli subscale of the Post-traumatic stress disorder scale of the Schedule for affective disorders and schizophrenia for school—age children present and lifetime version) - 12 month follow-up: There was a medium to large difference between participants randomised to the individual therapy plus caregiver support group and those randomised to the group psychotherapy plus caregiver support group (significance value not reported).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Overaschel's Re-experience of the traumatic event subscale, and Persistent avoidance of stimuli subscale of the Post-traumatic stress disorder scale (an extension of the Schedule for affective disorders and schizophrenia for school-age children present and lifetime version, K-SADS-PL) - used in Trowell et al. 2002. NB. The scale used to measure post-traumatic stress disorder in Deblinger et al., 2001 is not clear. No details are provided in relation to the authors of these scales or whether they had established reliability and validity. • Caregiver/parent health and wellbeing outcomes - This review did not record whether the included studies measured caregiver or parent health and wellbeing. • Satisfaction with services - This review did not record whether the included studies measured satisfaction with services. • Service outcomes - This review did not record whether the included studies measured service outcomes. **Follow-up:** Not consistently reported, however the authors note that Deblinger et al., 2001 only had a 3 month follow-up period, and that King		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	et al. 2000 had a 'brief' follow-up period. **Costs?** No. Costs and resource use information are not reported.		

7. Macdonald G, Higgins J, Ramchandani P et al. (2012) Cognitive-behavioural interventions for children who have been sexually abused: A systematic review. Campbell Systematic Reviews 2012: 14: 111

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: NB. This study is an updated version of a 2006 review and reports on those studies which were included in the earlier review. The objective of the review is to '… assess the efficacy of cognitive-behavioural approaches (CBT) in addressing the immediate and longer-term sequelae of sexual abuse on children and young people up to 18 years of age' (p6). **Methodology:** Systematic review of randomised and quasi-randomised controlled trials.	**Participants:** • Children and young people. Children up to the age of 18 who have experienced sexual abuse. • Caregivers and families – A number of the reviewed studies included non-offending parents (Celano et al. 1996; Cohen et al. 1996/Cohen and Mannarino, 1997/Cohen and Mannarino, 1996; Cohen et al. 2004/Cohen et al. 2006; Deblinger et al. 2001). **Sample characteristics -** • Age: The review protocol states that studies focusing on children and adolescents up to the age of 18 were to be included. The ages of participant children in the included studies ranged from two to 17. No details on the ages of participating parents are included. Berliner and Saunders, 1996 – 4–13 years. Burke, 1988 - 8–13 years.	**Statistical data:** The review identified 10 studies, giving a total sample of 847 participants (exact numbers of children and non-offending parents are not specified by the authors). Nine of the 10 studies included were conducted in the USA, and the tenth was conducted in Australia. NB. The authors do not report whether higher or lower scores on each scale represent an improvement. **Quality of parenting and parent-child relationships -** Parental belief and support of the child (measured using the Parental Support Questionnaire used in Cohen et al. 2004, and the Parents Reaction to Incest Disclosure Scale used in Celano et al. 1996. Total scores on these scales are not reported) - Short-term parental belief of child (immediately after treatment): Cognitive behavioural therapy had a small but significant effect; evaluated in two studies, 211 participants, standardised mean difference (IV, random, 95% CI), 0.30 [0.03, 0.57]. Intermediate term parental belief of child (three to 6 months after treatment: Cognitive behavioural therapy	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** ++ **Overall assessment of validity:** ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Country: Nine out of 10 of the trials reported by the included studies were conducted in the USA. The tenth (King et al. 2000) was conducted in Australia. The review was carried out by researchers in Northern Ireland. **Source of funding government:** • Government - Northern Ireland Research and Development. • Other - UK Nordic Campbell Center, Denmark.	Celano et al., 1996 – 8–13 years. Cohen et al., 1996 – 3–6 years. Cohen et al., 1998 – 7–15 years. Cohen et al., 2004 – 8–14 years. Deblinger et al., 1996 – 7–13 years. Deblinger et al., 2001 – 2–8 years. Dominguez, 2001 – 6–17 years. King et al., 2000 – 5–17 years. • Sex: Two studies included a nonoffending mother, or female caregiver as participants (Celano et al. 1996; Deblinger et al. 2001). Cohen et al. 1996; and Cohen et al. 2004 also included non-offending caregivers as participants but their gender is not reported. Berliner and Saunders 1996 – Male 11%. Female 89%. Burke 1988 – Female 100%. Celano et al. 1996 – Female 100%. Cohen et al. 1996 – Male 42%. Female 58%. (Treatment completers). Cohen et al. 1998 – Male 31%. Female 69%. (Treatment completers) Cohen et al. 2004 – Male 21%. Female 79%. Deblinger et al. 1996 – Male 17%. Female 83%. Deblinger et al. 2001 – Male 39%. Female 61%. (Treatment completers). Dominguez 2001 – Male 24%. Female 76%. King et al. 2000 – Male 31%. Female 69%.	had a small non-significant effect; evaluated in one study, 243 participants, standardised mean difference (IV, random, 95% CI), -0.32 [-0.65, 0.01]. Long term parental belief of child (at least one year): Cognitive behavioural therapy had a very small non-significant effect; evaluated in one study, 146 participants, standardised mean difference (IV, random, 95% CI), -0.10 [-0.43, 0.23]. Parental attributions (measured using the Parental Attributions Score. Total scores not reported.) - Self blame: Cognitive behavioural therapy had a large non-significant effect; evaluated in one study, 30 participants, mean difference (IV, random, 95% CI), -0.80 [-4.03, 2.43]. Child blame: Cognitive behavioural therapy had a large non-significant effect; evaluated in 1 study, 30 participants, mean difference (IV, random, 95% CI), -1.20 [-4.47, 2.07]. Perpetrator blame: Cognitive behavioural therapy had a medium non-significant effect; evaluated in 1 study, 30 participants, mean difference (IV, random, 95% CI), -0.60 [-2.62, 1.42]. Negative impact: Cognitive behavioural therapy had a large non-significant effect; evaluated in 1 study, 30 participants, mean difference (IV, random, 95% CI), -1.90 [-4.67, 0.87]. Parenting skills (measured using the measured using the Parenting Practices Questionnaire. Total scores not reported.) -	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Ethnicity: The majority of studies included participants who were either Caucasian, African American or Hispanic, however ethnicity is not reported at all for two studies (Burke 1998; King 2000). Berliner and Saunders, 1996 – Caucasian 74%, African American 12%, Hispanic 6%, 'other' 8%. Burke 1988 – Not reported by study. Celano et al. 1996 – Caucasian 22%, African American 75%, Hispanic 3%. Cohen et al. 1996 – Caucasian 54%, African American 42%, 'other' 4%. Cohen et al. 1998/Cohen et al., 2005 – Caucasian 59%, African American 37%, Hispanic 2%, Biracial 2%. Cohen et al. 2004 – White 60%, African American 28%, Biracial 7%, 'other' 1%. Deblinger et al. 1996 – Caucasian 72%, African American 20%, Hispanic 6%, 'other' 2%. Deblinger et al. 2001 – White 64%, Black 21%, Hispanic 2%, 'other ethnic origins' 14%. Dominguez 2001 – Caucasian 48%, African American 8%, Hispanic 40%, 'other' 4%. King et al. 2000 – Not reported by study.	Short-term parenting skills (immediately after treatment): Cognitive behavioural therapy had a large significant effect; evaluated in three studies, 278 participants, mean difference (IV, random, 95% CI), 3.86 [0.47, 7.26]. Intermediate term parenting skills (three to six months after treatment: Cognitive behavioural therapy had a large non-significant effect; evaluated in 3 studies, 231 participants, mean difference (IV, random, 95% CI), 2.36 [-1.55, 6.28]. Long term parenting skills (at least one year): Cognitive behavioural therapy had a large non-significant effect; evaluated in 2 studies, 193 participants, mean difference (IV, random, 95% CI), -0.89 [-4.89, 3.11]. Parent's emotional reactions (measured using the Parents' Emotional Reactions Questionnaire, follow-up point not clear) – Total score - Cognitive behavioural therapy had a large significant effect; evaluated in 2 studies, 558 participants, mean difference (IV, 95% CI), -5.17 [-7.17, -3.17]. Short-term parental emotional reactions (immediately after treatment): Cognitive behavioural therapy had a large significant effect; evaluated in 2 studies, 223 participants, mean difference (IV, 95% CI), -6.95 [-10.11, -3.80]. Intermediate term parental emotional reactions (three to six months after treatment: Cognitive behavioural therapy had a large non-significant effect; evaluated	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Religion/belief: Not reported for any of the included studies. • Disability: Not reported for any of the included studies. • Long term health condition: Not reported for any of the included studies. • Socioeconomic position: Not reported for any of the included studies, however the authors note that Cohen, et al. 1996, found that participants of a lower socioeconomic status were significantly more likely to complete treatment. • Type of abuse: All children had experienced sexual abuse. Perpetrators included family members (e.g. siblings, biological parents, grandparents, uncles, etc.), step-parents and partners of parents (usually mothers), and strangers/non-relatives. Severity and extent are not always reported but Berliner 1996; Cohen 1996; Cohen 1998; Deblinger 1996; Dominguez 2001; King 2000; Deblinger 2001 all report that some of the sample had experience of multiple/repeated sexual abuse. Details on any other types of maltreatment or abuse these children may have experienced is not provided with the exception of Berliner 1996; Burke 1988; Cohen 1998; Dominguez	in 2 studies, 187 participants, mean difference (IV, 95% CI), -3.46 [-6.98, 0.06]. Long term parental emotional reactions (at least one year): Cognitive behavioural therapy had a large significant effect; evaluated in 1 studies, 148 participants, mean difference (IV, 95% CI), -4.56 [-8.37, -0.75]. **Children and young people's health and wellbeing outcomes -** Depression (measured using the Children's Depression Inventory. Total scores not reported.) - Short-term depression (immediately after treatment): Cognitive behavioural therapy had a large non-significant effect; evaluated in 5 studies, 421 participants, mean difference (IV, random, 95% CI), -1.92 [-4.24, 0.40], I2 = 53%; p value for heterogeneity = 0.08. Intermediate term depression (three to six months after treatment: Cognitive behavioural therapy had a large significant effect; evaluated in 4 studies, 286 participants, mean difference (IV, random, 95% CI), -1.84 [-3.41, -0.27]. Long term depression (at least 1 year): Cognitive behavioural therapy had a large non-significant effect; evaluated in 4 studies, 301 participants, mean difference (IV, random, 95% CI), -1.19 [-2.70, 0.32]. Post-traumatic stress disorder (measured using a range of scales – not specified by review authors. Total scores not reported.)	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	2001, which also report details of physical abuse and/or risk of injury. • Looked after or adopted status: Not reported for any of the included studies. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported for any of the included studies. **Sample size:** • Sample size - The 10 included studies of the review gave a total sample size of 847 participants (it is assumed that this is the total number of children and young people who participated. The review authors do not provide details on the total number of non-offending parents who participated). The review authors record how many participants their analysis is based on for each outcome measure however the sample size for each study was: Berliner and Saunders 1996 – n=154 randomised. The number of participants assigned to each condition is reported as unclear by the review authors. Burke 1988 – n=25. Control n=12; intervention n=13. Celano et al., 1996 – n=49 randomised. Control n=24; intervention n=25. Also included non-offending female caretakers.	Short-term post-traumatic stress disorder (immediately after treatment): Cognitive behavioural therapy had a small to medium significant effect; evaluated in 6 studies, 442 participants, standardised mean difference (IV, random, 95% CI), -0.44 [-0.73, -0.16], I2 = 46%; P value for heterogeneity = 0.10. Intermediate term post-traumatic stress disorder (three to six months after treatment: Cognitive behavioural therapy had a small to medium significant effect; evaluated in 5 studies, 327 participants, standardised mean difference (IV, random, 95% CI), -0.39 [-0.74, -0.04]. Long term post-traumatic stress disorder (at least 1 year): Cognitive behavioural therapy had a small to medium significant effect; evaluated in 3 studies, 246 participants, standardised mean difference (IV, random, 95% CI), -0.38 [-0.65, -0.11]. Anxiety (Scale not specified by review authors. Total scores not reported.) Short-term anxiety (immediately after treatment): Cognitive behavioural therapy had a small significant effect; evaluated in 5 studies, 434 participants, standardised mean difference (IV, fixed, 95% CI), -0.23 [-0.42, -0.03] I2 = 0%; P value for heterogeneity = 0.84. Intermediate term anxiety (three to six months after treatment): Cognitive behavioural therapy had a small to medium significant effect; evaluated in 4 studies, 296 participants, standardised mean difference (IV, fixed, 95% CI), -0.38 [-0.61, -0.14].	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Cohen et al. 1996 – n=86 randomised. The number of participants assigned to each condition is reported as unclear by the review authors. Also included non-offending parents. Cohen et al. 1998 – n=82 randomised. Control n=41; intervention n=41. Cohen et al. 2004 – N=229 randomised. Control n=115; intervention n=114. Also included n=189 parents/caretakers. Deblinger et al. 1996 – n=100 randomised. Control n=25; intervention 1 n=25; intervention 2 n=25; intervention 3 n=25. Deblinger et al., 2001 – n=54 randomised. The number of participants assigned to each condition is reported as unclear by the review authors. Also included non-offending mothers. Dominguez 2001 – n=32 randomised. Control n=10; intervention n=22. King et al. 2000 – n=36 randomised. Control n=12; intervention 1 n=12; intervention 2 n=12. • Systematic reviews: Number of studies – 10 trials reported in 15 studies. **Intervention:**	Long term anxiety (at least 1 year): Cognitive behavioural therapy had a small significant effect; evaluated in 4 studies, 278 participants, standardised mean difference (IV, fixed, 95% CI), -0.28 [-0.52, -0.04]. Child sexualised behaviour (measured using the Child Sexual Behavior Inventory. Total scores not reported.) - Short-term child sexualised behaviour (immediately after treatment): Cognitive behavioural therapy had a medium non-significant effect; evaluated in 5 studies, 451 participants, mean difference (IV, random, 95% CI), -0.65 [-3.53, 2.24], I2 = 67%, p value for heterogeneity 0.02. Intermediate term child sexualised behaviour (3 to 6 months after treatment): Cognitive behavioural therapy had a small to medium non-significant effect; evaluated in 3 studies, 133 participants, mean difference (IV, random, 95% CI), -0.46 [-5.68, 4.76]. Long term child sexualised behaviour (at least 1 year): Cognitive behavioural therapy had a large non-significant effect; evaluated in 3 studies, 161 participants, mean difference (IV, random, 95% CI), -1.61 [-5.72, 2.49]. Child externalising behaviour (measured using the Child Behaviour Checklist) - Short-term child externalising behaviour (immediately after treatment): Cognitive behavioural therapy had a very small non-significant effect; evaluated in 7 studies, 537 participants, standardised mean difference (IV, random, 95% CI), -0.12 [-0.40, 0.17].	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Intervention category - Other – Cognitive behavioural therapy. • Describe intervention - An overall description of cognitive behavioural therapy is provided in the review. The review authors describe cognitive behavioural treatment as deriving from 4 theories of learning: 'respondent conditioning (associative learning); operant conditioning (the effect of the environment on patterns of behaviour); observational learning (learning by imitation); and cognitive learning (the impact of thought patterns on feelings and behaviour)' (p13). The authors state that children who experience sexual abuse may experience 'psychobiological changes that contribute to the development and maintenance of post-traumatic stress symptoms' (p14). These include affective, behavioural, cognitive, complex PTSD and psychobiological trauma symptoms. The authors state that cognitive behavioural therapy is designed to address these symptoms through a range of techniques. Emotional distress: Children helped to cope with emotional distress, for example through learning about relaxation and emotional expression skills.	Intermediate term child externalising behaviour (3 to 6 months after treatment): Cognitive behavioural therapy had a very small non-significant effect; evaluated in four studies, 175 participants, standardised mean difference (IV, random, 95% CI), -0.11 [-0.42, 0.21]. Long term child externalising behaviour (at least one year): Cognitive behavioural therapy had a very small non-significant effect; evaluated in five studies, 355 participants, standardised mean difference (IV, random, 95% CI), 0.05 [-0.16, 0.27]. **Narrative findings - effectiveness -** Quality of the evidence: The authors report that the quality of the included studies was not of a high standard and that there were a range of methodological weaknesses, e.g. in relation to randomisation or blinding of investigators as well as a generally poor level of detail provided on potential sources of bias. A particular source of concern seems to be that only three of the included studies stated that analysis was conducted on an intent-to-treat basis. **Quality of parenting and parent-child relationships -** Parental belief and support of the child (measured using the Parental Support Questionnaire used in Cohen et al. 2004, and the Parents Reaction to Incest Disclosure Scale used in Celano et al. 1996. Total scores on these scales are not reported) - Cognitive behavioural therapy was found to have a very small but significant effect on parental belief and support of the child immediately after treatment; however in the intermediate term (3 to 6 months after	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Anxiety: Children taught to recognise the signs of anxiety, and how to replace maladaptive responses to anxiety with adaptive ones. Behaviour problems: Parents supported to understand the impact of sexual abuse on children's behaviour, and how this is shaped or maintained by consequences. • Delivered by - Not reported for any of the included studies. • Delivered to - Children and adolescents between the ages of 2 and 17. Two interventions appear to have been delivered to children on their own (Berliner 1996; Burke 1998), whilst the majority also involved a non-offending parent to some extent, either in joint sessions or in parallel sessions. • Some interventions appear to have been delivered in group format (e.g. Burke 1988; Celano 1996). • Duration, frequency, intensity, etc. - The number of sessions is not always reported. Where this detail is provided, interventions generally lasted from between 8 to 12 sessions, however some provided as few as 6 (Burke, 1988) whilst others provided as many as 20 (King 2000). Sessions typically seem to be delivered on a weekly basis and are at least 45 minutes in duration.	treatment) and in the longer term (at least one year) the effects were very small to small and non-significant. Short-term parental belief of child (immediately after treatment): Cognitive behavioural therapy (evaluated in two studies, giving a combined sample of 211 participants) had a small non-significant effect. Intermediate term parental belief of child (three to six months after treatment): Cognitive behavioural therapy (evaluated in one study, with a sample of 243 participants) had a small non-significant effect. Long term parental belief of child (at least 1 year): Cognitive behavioural therapy (evaluated in 1 study, with a sample of 146 participants) had a very small non-significant effect. Parental attributions (measured using the Parental Attributions Score. Total scores not reported.) - Cognitive behavioural therapy was found to have medium or large, but non-significant impact on parental self blame, child blame, perpetrator blame and negative impact. Self blame: Cognitive behavioural therapy (evaluated in 1 study, with a sample of 30 participants) had a large non-significant effect. Child blame: Cognitive behavioural therapy (evaluated in 1 study, with a sample of 30 participants) had a large non-significant effect.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Key components and objectives of intervention – To address the sequelae of sexual abuse by focusing on the meaning of events and identifying maladaptive cognitions and misattributions. The authors note that the majority of included studies are described as trauma-focused '… meaning simply that the adverse consequences of child sexual abuse are conceptualised as the consequences of trauma, which is reflected in the structure and the content of treatment' (p26). • Content/session titles - Not reported consistently. The titles of programmes which are reported include Recovering from Abuse Program (Celano et al. 1996); Cognitive-behavioural therapy for sexually abused children (Cohen et al. 1996); and Sexual Abuse Specific Cognitive behavioural therapy (Cohen et al. 1998). Berliner and Saunders 1996 – As for the control group - 10 sessions covering – '…getting acquainted and establishing ground rules; feelings; family and friends (2 sessions); disclosure impact, self-esteem and sexual abuse; body awareness and sexuality (2 sessions), and prevention and termination' (p30). For the intervention	Perpetrator blame: Cognitive behavioural therapy (evaluated in one study, with a sample of 30 participants) had a medium non-significant effect. Negative impact: Cognitive behavioural therapy (evaluated in one study, with a sample of 30 participants) had a large non-significant effect. Parenting skills (measured using the measured using the Parenting Practices Questionnaire. Total scores not reported.) - Cognitive behavioural therapy had a large significant effect on short-term parenting skills, but this was not significant, although effect sizes were still large, in the intermediate and long term. Short-term parenting skills (immediately after treatment): Cognitive behavioural therapy (evaluated in three studies, giving a combined sample of 278 participants) had a large significant effect. Intermediate term parenting skills (3 to 6 months after treatment: Cognitive behavioural therapy (evaluated in three studies, giving a combined sample of 231 participants) had a large non-significant effect. Long term parenting skills (at least 1 year): Cognitive behavioural therapy (evaluated in two studies, giving a combined sample of 193 participants) had a large non-significant effect. Parent's emotional reactions (measured using the Parents' Emotional Reactions Questionnaire, follow-up point not clear) –	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	group, the session on feelings specifically focused on fear; 1 of the family and friends sessions was replaced with a Stress Inoculation Therapy session (no further details provided); 2 sessions on gradual exposure were provided; and Stress Inoculation Therapy principles were applied to the sessions on impact of disclosure and self-esteem. Burke 1988 – 6 group sessions (group format) focusing on 'good and bad touching', anxiety and relaxation techniques, imaginal exposure and identifying feelings of depression, identifying pleasurable events and being able to engage in them; development of strategies to deal with 'bad touching' and the anxiety this creates. Celano et al. 1996 – Recovering from Abuse Program (eight group sessions) which focus on maladaptive affects, beliefs, and behaviour relating to betrayal; powerlessness; self-blame and stigmatisation; traumatic sexualisation. Cohen et al. 1996/Cohen and Mannarino 1997/Cohen and Mannarino 1996 – Cognitive behavioural therapy for sexually abused children. No further details provided.	Cognitive behavioural therapy had a large significant effect on parents' emotional reactions (as measured by the parent's emotional reactions questionnaire) in terms of total scores, and in the short and long terms. In the intermediate term, the effect size was large but not significant. Total score - Cognitive behavioural therapy (evaluated in two studies, giving a combined sample of 558 participants) had a large significant effect. Short-term parental emotional reactions (immediately after treatment): Cognitive behavioural therapy (evaluated in two studies, giving a combined sample of 223 participants) had a large significant effect. Intermediate term parental emotional reactions (3 to 6 months after treatment: Cognitive behavioural therapy (evaluated in two studies, giving a combined sample of 187 participants) had a large non-significant effect. Long term parental emotional reactions (at least 1 year): Cognitive behavioural therapy (evaluated in one study with a sample of 148 participants) had a large significant effect. **Children and young people's health and wellbeing outcomes -** Depression (measured using the Children's Depression Inventory. Total scores not reported.) - Cognitive behavioural therapy showed large effect sizes on depression in the short term (but this was not statistically significant), intermediate term (statistically significant) and long term (not statistically significant).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Cohen et al. 1998/Cohen et al. 2005 – Sexual Abuse Specific Cognitive behavioural therapy. Children's sessions focus on anxiety, behavioural problems and depression. Parental sessions focus on parental emotional distress and enabling the parent to manage behavior and provide emotional support. Cohen et al., 2004 – Trauma-focused cognitive behavioural therapy. Individual sessions provided to both parent and child and three joint child-parent sessions. Deblinger et al. 1996 – Evaluated three experimental conditions – an intervention delivered solely to children; an intervention delivered solely to parents; and an intervention delivered to both children and parents. The intervention delivered to children included body safety training, coping, education, gradual exposure, and modelling. The intervention delivered to parents aimed to enable mothers with the cognitive behavioural skills to respond to their child's avoidance and fear behaviours. The joint intervention individual child and parent sessions as well as joint sessions which aimed to enhance child-parent communication to enable therapeutic work to be continued in the home environment.	Short-term depression (immediately after treatment): Cognitive behavioural therapy (evaluated in five studies, giving a combined sample of 421 participants) had a large non-significant effect. Intermediate term depression (3 to 6 months after treatment: Cognitive behavioural therapy (evaluated in four studies, giving a combined sample of 286 participants) had a large significant effect. Long term depression (at least 1 year): Cognitive behavioural therapy (evaluated in four studies, giving a combined sample of 301 participants) had a large non-significant effect. Post-traumatic stress disorder (measured using a range of scales – not specified by review authors. Total scores not reported.) Cognitive behavioural therapy had small to medium effect, which was statistically significant, on post-traumatic stress disorder symptoms in the short, intermediate and long term. Short-term post-traumatic stress disorder (immediately after treatment): Cognitive behavioural therapy (evaluated in six studies, giving a combined sample of 442 participants) had a small to medium significant effect. Intermediate term post-traumatic stress disorder (3 to 6 months after treatment: Cognitive behavioural ther-	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Deblinger et al. 2001 – Manualised intervention. Parents receive 11 sessions of cognitive behavioural therapy aiming to enable parents to - deal with their own emotional reactions so that they can support their children; foster child-parent communication about the abusive experience; and manage behavioural problems displayed by the child. Children also receive 11 sessions aiming to - help the child to communicate their feelings and learn how to cope with these; identify appropriate and inappropriate touching; and to '… learn abuse response skills, using an interactive behavioural format' (p44). Dominguez 2001 – The main goal of treatment was to provide children with the skills to manage affective, behavioural and cognitive responses to the traumatic events. King et al. 2000 – Evaluated 2 experimental conditions. Child Cognitive-Behavioural Therapy which aims to enable children to overcome post-abuse distress and post-traumatic stress disorder symptoms. Family Cognitive Behavioural Therapy plus sessions for parents in behaviour management techniques and child-parent communication.	apy (evaluated in 5 studies, giving a combined sample of 327 participants) had a small to medium significant effect. Long term post-traumatic stress disorder (at least one year): Cognitive behavioural therapy (evaluated in three studies, giving a combined sample of 246 participants) had a small to medium significant effect. Anxiety (Scale not specified by review authors. Total scores not reported.) Cognitive behavioural therapy had a small, statistically significant effect on anxiety in the short, intermediate (small to medium effect size) and long term. Short-term anxiety (immediately after treatment): Cognitive behavioural therapy (evaluated in 5 studies, giving a combined sample of 434 participants) had a small significant effect. Intermediate term anxiety (3 to 6 months after treatment): Cognitive behavioural therapy (evaluated in four studies, giving a combined sample of 296 participants) had a small to medium significant effect. Long term anxiety (at least 1 year): Cognitive behavioural therapy (evaluated in four studies, giving a combined sample of 278 participants) had a small significant effect. Child sexualised behaviour (measured using the Child Sexual Behavior Inventory. Total scores not reported.)	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Location/place of delivery - Not reported for any of the included studies. • Describe comparison intervention - The review protocol specified that studies '… comparing CBT versus treatment as usual (referred to in the protocol as 'another intervention'), with or without placebo control, were eligible, as were studies comparing one intervention versus control' (p15). The authors report that although one study used a wait list control (Burke 1988) the other nine studies compared cognitive behavioural therapy to treatment as usual '… which was typically supportive unstructured psychotherapy' (p23). Berliner and Saunders 1996 – As for the intervention group - 10 sessions covering – '…getting acquainted and establishing ground rules; feelings; family and friends (2 sessions); disclosure impact, self-esteem and sexual abuse; body awareness and sexuality (2 sessions), and prevention and termination' (p30). Burke 1988 – Waitlist control. Celano et al. 1996 – Treatment as usual which the review authors' note was '… defined as supportive, unstructured psychotherapy that sexually abused children and their	Cognitive behavioural therapy did not have a statistically significant impact on child sexualised behaviour in the short, intermediate or long term. However, effect sizes were medium, small to medium and large respectively. Short-term child sexualised behaviour (immediately after treatment): Cognitive behavioural therapy (evaluated in 5 studies, giving a combined sample of 451 participants) had a medium non-significant effect. Intermediate term child sexualised behaviour (three to six months after treatment): Cognitive behavioural therapy (evaluated in three studies, giving a combined sample of 133 participants) had a small to medium non-significant effect. Long term child sexualised behaviour (at least one year): Cognitive behavioural therapy (evaluated in three studies, giving a combined sample of 161 participants) had a large non-significant effect. Child externalising behaviour (measured using the Child Behaviour Checklist) - Cognitive behavioural therapy did not have a statistically significant impact on child externalising behaviour, and very small effect sizes were observed, in the short, intermediate and long term. Short-term child externalising behaviour (immediately after treatment): Cognitive behavioural therapy (evaluated in seven studies, giving a combined sample of 537 participants) had a very small non-significant effect.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	mothers would normally receive at the clinic' (p34). Cohen et al. 1996 - Non-directive supportive therapy. Cohen et al. 1998 - Non-specific therapy which '… did not provide suggestions or directive advice, but encouraged exploration of alternative attributions, behaviours and feelings via nondirective interventions' (p38). Cohen et al. 2004 – Not reported. Deblinger et al. 1996 – Described as 'community control' which provides caregivers with information in relation to symptom patterns and encourages caregivers to access therapeutic care. Deblinger et al. 2001 – Manualised supportive group therapy (11 sessions) which aim to empower caregivers. The therapists running these sessions do not provide information which specifically relates to behaviour management, coping, or gradual exposure. As in the intervention group, children in the control group also receive 11 sessions aiming to - help the child to communicate their feelings and learn how to cope with these; identify appropriate and inappropriate touching; and to '… learn abuse response skills, using an interactive	Intermediate term child externalising behaviour (3 to 6 months after treatment): Cognitive behavioural therapy (evaluated in four studies, giving a combined sample of 175 participants) had a very small non-significant effect. Long term child externalising behaviour (at least 1 year): Cognitive behavioural therapy (evaluated in five studies, giving a combined sample of 355 participants) had a very small non-significant effect.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	behavioural format' (p44). However, therapists '… used a didactic format, presenting age appropriate information and personal safely using pictures, stories and activity age exercises' (p 44). Dominguez 2001 – Supportive treatment which aims to '… facilitate change via a combination of consciousness raising and corrective emotional experiences that occur in the context of a genuine, empathic relationship characterised by unconditional positive regard' (p46). King et al. 2000 – Waitlist control. • Delivered by - Not reported for any of the included studies. • Delivered to - Children and adolescents. • Duration, frequency, intensity, etc. - Not reported for any of the included studies. • Key components and objectives of intervention - Not reported for any of the included studies. • Content/session titles – Not reported for any of the included studies. • Location/place of delivery - Not reported for any of the included studies. **Outcomes measured:**		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Incidence of abuse and neglect - This review did not record whether the included studies measured incidence of abuse and neglect. • Risk of abuse and neglect - This review did not record whether the included studies measured risk of abuse and neglect. • Quality of parenting and parent-child relationships - The authors describe the outcome measures used by individual studies as 'parental skills and knowledge' which included behavioural management skills, parental emotional reactions, knowledge of child sexual abuse and possible consequences of this, 'belief in their child's story', and understanding of child behaviour and psychological problems. The outcome measures used by individual studies which are reported in the review are: Parental belief and support was measured using the Parental Support Questionnaire (revised, used in Cohen et al. 2004); and the Parents Reaction to Incest Disclosure Scale (used in Celano et al. 1996). NB The authors state that this measure was not used in their review on p53 however information the section of the findings table dealing with parental belief and support of the child suggests that		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	scores on this measure were combined with those on the Parental Support Questionnaire (revised) in the meta-analysis. Parental attributions were measured using the Parental Attributions Score (used in Celano et al. 1996). Parenting practices/skills were measured using the Parenting Practices Questionnaire (used in Deblinger et al. 1996; Deblinger et al. 2001; and Cohen et al. 2004). Parental emotional reaction was measured using the Parents' Emotional Reactions Questionnaire (used in Deblinger et al. 2001; Cohen et al. 2004). • Children and young people's health and wellbeing outcomes - The review protocol states that the primary outcomes that would be considered were children's psychological functioning (e.g. anxiety, depression and post-traumatic stress disorder); and children's behavioural problems such as externalising and sexualised behaviour. Secondary outcomes are summarised as 'future offending behaviours' such as sexual offending, delinquency and criminal offending; and outcomes relating to the quality of parenting and parent-child relationships which the authors categorise as 'parental skills and knowledge'		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	(included behavioural management skills, parental emotional reactions, knowledge of child sexual abuse and possible consequences of this); 'belief in their child's story'; and understanding of child behaviour and psychological problems. NB. The review does not report on 'future offending behaviours' and it is therefore not clear if any of the included studies report on this type of outcome. Child depression was measured using: The Children's Depression Inventory (used in seven studies - Berliner 1996; Burke 1988; Cohen 1998; Cohen 2004; Deblinger 1996; Dominguez 2001; King 2000) however the meta-analysis only appears to use data from 5 studies and it is unclear which these are. Child post-traumatic stress disorder was measured using: The Children's Impact of Traumatic Events Scales-Revised (used in Celano et al. 1996). The Post-traumatic stress disorder subscale of the Kiddie-Schedule for Affective Disorders and Schizophrenia - Epidemiologic version (used in Deblinger, et al. 2001). The post-traumatic stress disorder supplement (re-experiencing) of the Schedule for Affective Disorders and Schizophrenia for school-age		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	children - Present and Lifetime version (used in Cohen et al. 2004). The Trauma Symptom Checklist for Children - Post-traumatic stress disorder (used in Cohen et al. 1998). The re-experiencing subscale of the post-traumatic stress disorder section of the Anxiety Disorders Interview Schedule DSM IV, child version (used in King et al. 2000). Child anxiety was measured using: The Revised Children's Manifest Anxiety Scale (total scores, used in Berliner et al. 1996; and King. et al, 2000. The State/Trait Anxiety Inventory for Children (state scale scores only, used in Cohen et al. 1998; Deblinger et al. 1996; and Cohen et al. 2004). Sexualised behaviour was measured using: The Child Sexual Behavior Inventory (used in Berliner et al. 1996; Cohen et al. 1996; Cohen et al. 1998; Deblinger et al. 2001; and Cohen et al. 2004). Child externalising behaviour was measured using: The externalising behaviour scale of the Child Behaviour Checklist (parent report, used in Berliner et al. 1996; Celano et al. 1996; Cohen et al. 1996; Cohen et al. 1998; Deblinger et al. 1996;		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	King et al. 2000; Deblinger et al. 2001; and Cohen et al. 2004). • Caregiver/parent health and wellbeing outcomes - This review did not record whether the included studies measured caregiver or parent health and wellbeing. • Satisfaction with services - This review did not record whether the included studies measured satisfaction with services. • Service outcomes - This review did not record whether the included studies measured service outcomes. **Follow-up:** NB. Although follow-up periods vary between studies, the authors have categorised these into short-term, intermediate, or long-term. Berliner et al. 1996 – Post-treatment, 12 months, and 24 months. Burke et al. 1988 – Not reported. Celano et al. 1996 – Post-treatment only. Cohen et al. 1996 – Post-treatment, six months and 12 months. Cohen et al. 1998 - Post-treatment, six months and 12 months. Cohen et al. 2004 – Post-treatment only. Deblinger et al. 1996 – Post-treatment, 3 months, 6 months, 12 months, and 24 months.		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Deblinger et al. 2001 – At around the 11 week mark of the intervention period, and three months post-treatment. Dominguez et al. 2001 – Post-treatment only. King et al., 2000 – Post-treatment, and 12 weeks. **Costs?** No. Cost and resource use information are not provided.		

8. Parker B and Turner W (2013) Psychoanalytic/psychodynamic psychotherapy for children and adolescents who have been sexually abused: A systematic review. Campbell Collaboration 9(13)

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To '…assess the effectiveness of psychoanalytic/psychodynamic psychotherapeutic approaches in treating the effects of sexual abuse (psychologically and in terms of behaviour and social functioning) in children and adolescents' (p13). **Methodology:** Systematic review of randomised and quasi-randomised controlled	**Participants:** Children and young people. The review aimed to include studies in which the sample was comprised of children and adolescents up to the age of 18 years with experience of sexual abuse. **Sample characteristics:** N/A The review did not identify any studies which were eligible for inclusion. **Sample size:** N/A The review did not identify any studies which were eligible for inclusion. **Intervention:**	**Findings:** The review did not identify any studies which were eligible for inclusion.	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** ++ **Overall assessment of validity:** ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
trials using a no treatment control or wait list control (studies with an active comparison group were excluded). **Country:** The review was conducted by authors based in England. The review did not identify any studies which were eligible for inclusion. **Source of funding:** Other - Centre for Gender Violence Research, School for Policy Studies, University of Bristol, UK South West London and St Georges Mental Health NHS Trust, UK.	• Intervention category - Other – The review aimed to evaluate interventions described as psychoanalytic/psychodynamic psychotherapy. • Describe intervention – The review authors describe psychoanalytic and psychodynamic psychotherapy as umbrella terms which are often used interchangeably and cover a range of approaches and techniques such as child or adolescent psychotherapy; child analysis or psychoanalysis; Freudian, Jungian, or Kleinian therapy; object relations based therapy; etc. Despite this variation in usage the authors note that the defining features of these approaches is the objective of forming a therapeutic relationship and through this exploring (through discussion or play) how earlier events can impact upon current behaviours, feelings, and relationships. The goal of this process is to enable the individual to become aware of previously 'unconscious' difficulties. The review authors note that psychoanalytic/psychodynamic therapy can be provided both individually or to families or larger groups, and can be of varying intensity with treatment being provided for a few brief sessions, or being delivered over the course of		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	years. The authors note that the intervention could be of any length and could also include separate work with caregivers or parents or sessions which were delivered to both the child and their caregiver of parent. **Comparison intervention:** The review aimed to evaluate psychoanalytic/psychodynamic psychotherapy by comparison to no treatment or wait list controls and studies with an active comparison group (such as cognitive behavioural therapy) were excluded. **Outcomes measured:** The authors did not identify any studies that were eligible for inclusion. The outcomes of interest to the review related to children's and young people's health and wellbeing (i.e. aggression and conduct problems, depression, functioning, post-traumatic stress disorder, psychiatric symptoms, self-harm, sexualised behaviour and suicide). **Follow-up:** N/A The review did not identify any studies which were eligible for inclusion. **Costs?** No. The authors do not specify whether they intended to record cost or resource use information.		

9. Shirk SR, DePrince AP, Crisostomo PS et al. (2014) Cognitive behavioral therapy for depressed adolescents exposed to interpersonal trauma: An initial effectiveness trial. Psychotherapy 51: 167–79

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim To assess the feasibility, acceptability and initial impact of a modified CBT protocol (based on 1 previously found to be effective, the Adolescent Mood Project protocol) for adolescent depression (m-CBT) designed to treat adolescents with history of interpersonal trauma. **Methodology** RCT including cluster. Adolescents referred to 2 American outpatient child and adolescent clinics at an urban community mental health centre were randomised to receive 12 sessions of m-CBT treatment or usual care, following a pre-treatment assessment, using a stratified randomisation procedure (using gender variable). Post-treatment	**Participants** Children and young people. Adolescents referred to 2 outpatient community mental health clinics, who were not: (1) receiving concurrent psychological treatment for depression, (2) to have attempted suicide within 3 months before intake, (3) engaged in self-injurious behavior that required hospitalisation or emergency room treatment within the past 3 months, (4) meeting diagnostic criteria for bipolar disorder and/or substance dependence disorder, (5) presenting with psychotic symptoms or intellectual deficit (i.e. estimated IQ below 70). **Sample characteristics:** • Age – Mean age 15.48, range 13-17, SD=1.53 m-CBT mean age 15.25 (1.52) UC mean age 15.69 (1.55). • Sex – n=36 female, n=7 male. • Ethnicity – 49% non-Hispanic Caucasian. Of the other 51% - 33% Hispanic, 38% African American. m-CBT 55% ethnic minority UC 47.83 % ethnic minority. • Religion/belief – Not reported. • Disability – Not reported.	**Effect sizes** **Children and young people's health and wellbeing** Beck Depression Inventory A repeated measures linear mixed-methods model analysed BDI-II scores looking at treatment condition and time as factors. Only female (n=36) participants were analysed due to the small male sample (n=7) who had had no observations in Sessions 8 and 12 in the m-CBT group. No significant effects were found, either for which clinic was attended (F(1, 45)=2.65, p=.12); treatment condition (F (1, 54)=.09, p=.78) or the interaction between time and condition (F(5,128)=1.80, p=.12). BDI-II scores decreased in both groups up to session 12 and then rose again at posttreatment assessment. When male data was included, using a last-observation-carried-forward analysis model for missing BDI-II scores there was still no significant effect seen for either condition (F(1,42)=.06, p=.81) or time-by-condition (F(1,42) =1.76, p=.19). The effect size for between-group differences at post-treatment slightly favoured UC (d=.16, t(42)=1.27, p=.18) but this was not reliable. The effect for time was significant, (F, 41)=27.20, p<.001) reflecting symptom reduction in both groups - on average BDI-II scores dropped by 12.83 points in the UC condition and 8.50 in the m-CBT group. Fisher's exact test was used to assess the effect of treatment on depression diagnosis at posttreatment -	**Overall assessment of internal validity** - **Overall assessment of external validity** ++ Despite US context both m-CBT and usual care (client-centred or psychodynamic interventions) are similar to the UK context. **Overall validity score** - Key limitations of the study include relatively small sample size, no male participants completed the study. Unclear what the analysed sample size was, due to exclusion of male participants.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
assessment was conducted at 16 weeks after treatment began but participants were able to carry on treatment after assessment. All treatment sessions were audio-recorded. The measures used for diagnosis, screening or assessment were: Kiddle-Schedule for Affective Disorders and Schizophrenia-Present and Lifetime version (K-SADS-LS) - to diagnose depressive , posttraumatic stress, substance abuse and dependence disorders and to screen for bipolar disorder, suicide attempts and psychotic symptoms ; an abbreviated version of the Trauma Experiences Screening Inventory-Child Version (TESI-C)- to determining presence of prior trauma; the Beck Depression Inventory (BDI-II) - to assess severity of depressive	• Long term health condition – Eligible participants met criteria for depressive disorder (using K-SADS): Major depressive disorder n=35 Dysthymic disorder n=3 Depressive disorder not otherwise specified n=5. • Sexual orientation – Not reported. • Socioeconomic position – Not reported. • Type of abuse – 'All eligible adolescents reported at least one incident of physical abuse (49%), witnessing family violence (58%), sexual abuse (67%), and verbal/emotional abuse (47%) in response to a highly structured screening interview. A majority of the sample reported more than one type of interpersonal trauma throughout their lifetime: 1 type (23%); 2 types (28%); 3 or more types (49%).'(p.12) % Sexually abused 66.66% m-CBT group 68.18 UC group % Physically abused 60.00 m-CBT group 40.90 UC group % Emotionally abused 55.00 m-CBT group 42.80 UC group % Witnessed domestic violence 65.00 m-CBT group 54.50 UC group Number of trauma types 2.80 (1.20) m-CBT group 2.45 (1.26) UC group. • Looked after or adopted status – Not reported.	it found that there was no significant difference between groups (p=.92) - diagnostic remission rates for the full sample was 48% for UC and 50% for m-CBT. **Satisfaction with services** Multivariate analysis was carried out on CSQ and TEI data from the n=36 adolescents who completed the posttreatment assessment. It found that m-CBT and UC participants did not differ on client satisfaction (CSQ) or treatment acceptability (TEI) but found high scores for both in both conditions. Difference F(2, 31)=.02, p<.98.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
symptoms at pre- and posttreatment assessments as well as after Sessions 1, 4, 8 and 12; the Child Behaviour Checklist (CBCL) to evaluate 'collateral disruptive behaviour problems' (p171), completed by the adolescent's mother or guardian pre-and post-treatment'; Similarities (and Block Design if adolescent scored <7 due to English not being their first language) subtest of the Weschler Adult Intelligence Scale-IV or the Weschler Intelligence Scale for Children-IV, to measure verbal IQ and conceptual ability at preassessment; the Therapy Process Observational Coding System- Strategies Scale (TPOCS-S) to evaluate therapist interventions in the UC condition, with sessions from early and later phases randomly picked for coding on	• Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size** Comparison numbers - UC group n=23 Intervention numbers - m-CBT group n=20 Sample size - n=43 n=20 m-CBT group n=23 UC group **Intervention category** Other Modified CBT **Intervention** Describe intervention Individual therapy for depressed adolescents with a history of interpersonal trauma (physical, sexual and/or emotional abuse) delivered in 2 outpatient clinics over 12 weeks with weekly sessions guided by a manual. Treatment, either for intervention or usual care was not time-limited; n=8 adolescents continued to receive the m-CBT treatment after the 16 week posttreatment assessment, n=10 the usual treatment. Therapists able to refer adolescents for medication at any time, and were informed of each adolescent's diagnosis and history before commencing treatment. Delivered by		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
use of psychodynamic, cognitive, behavioural, client-centred and family systems; the Adherence Checklist for m-CBT, adapted from the adherence checklist for the original protocol, a Client Satisfaction Questionnaire (CSQ) (scores ranging 8–32 with higher scores indicating higher satisfaction) and the Treatment Evaluation Inventory (TEI) to assess treatment acceptability at posttreatment (15 items rated for fairness, appropriate and reasonable nature). At the 16 week posttreatment assessment participants completed the K-SADS depression and PTSD modules, the BDI-II, the CSQ and TEI with an independent evaluator. After testing for group comparability, missing data comparisons, and therapist differences, treatment effects were	2 clinic-based therapists, one male (Doctoral-level psychologist with 28 years clinical experience), 1 female (masters level therapist with 10 years of experience), who volunteered to deliver the m-CBT therapy sessions. They completed a 1 day workshop lead by an expert in mindfulness-based interventions (Elizabeth Roehmer) and the m-CBT developers (Anne P.de Prince and Stephen R. Shirk). It covered basic CBT principles and components of m-CBT (including practice with mindfulness exercises). Each therapist completed a practice case under supervision before the start of the clinical trial, and received weekly supervision from Anne P. DePrince throughout. Delivered to n=20 adolescents, mean age 15.25, who had been referred for outpatient treatment at 2 clinics at an urban mental health centre and met diagnostic criteria for a depressive disorder (details above). Duration, frequency, intensity, etc. m-CBT protocol is delivered through 12 weekly sessions of individual therapy (except for the first session where parents were invited to attend). Key components and objectives of intervention m-CBT has modules (based on the original protocol) looking at mood and		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
analysed for the primary outcome variables of feasibility, acceptability and impact (on depression outcomes), with an intention to treat sample. **Country** Not UK. US - Rocky Mountain region.	cognition; mood and activities; mood and interpersonal relationships. It combines CBT elements: mood monitoring, cognitive restructuring, relaxation exercises, activity scheduling and interpersonal problem-solving with mindfulness exercises and applications across all sessions. It encourages and emphasises the need to observe, describe and tolerate trauma-related emotions and cognitions in a non-judgemental way. This is designed to improve self-monitoring and enable people to live more in the present than in the past. Therapists are explicitly instructed to tackle the interpersonal trauma experience and cognitions relating to it. Content/session titles 1 Introduction to Therapy, Depression, and Mindfulness 2 Mindfulness: Learning to Observe 3 Mindfulness of Sights and Sounds: Learning to Describe 4 Mindfulness Now: Learning to Participate 5 Mindfulness of Thoughts 6 Noticing Thoughts: Hey, They're Not Facts! 7 What to Do with All those Fish in the Fish Tank? 8 Mindfulness of Trauma-Related Thoughts and Emotions 9 More on Mindfulness of Trauma-Related Thoughts and Emotions 10 Mindfulness of Relationships 11 Participating Mindfully in Relationships 12 Where Have We Come From and Where Do		

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	We Go From Here? Staying on Active Pilot!' Location/place of delivery At the outpatient clinics Describe comparison intervention 2 UC therapists, female doctoral-level psychologists (with 3 and 4 years of clinical experience, respectively) who volunteered to work with the participants randomised to UC used treatment strategies and procedures that they regularly used and believed to be effective -involving client-centred, psychodynamic and family interventions. Treatment did not follow a specific manual and were coded with TPOCS-S to describe the treatment strategies and how they differed from m-CBT. The treatment also took place across 12 weekly sessions, and participants were also able to continue after posttest assessment. **Outcomes measured** Children and young people's health and wellbeing outcomes Depressive symptoms based on BDI-II scores from pre-treatment, sessions 1, 4, 8, 12 and posttreatment assessment. Satisfaction with services Satisfaction, as measured by CSQ and TEI at posttreatment assessment.		

10. Trowell J, Kolvin I, Weeramanthri T et al. (2002) Psychotherapy for sexually abused girls: psychopathological outcome findings and patterns of change. The British journal of psychiatry: the journal of mental science 180: 234–47

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim 'To compare the relative efficacy of focused individual or group therapy in symptomatic sexually abused girls, and to monitor psychiatric symptoms for persistence or change' (p234). **Methodology** RCT including cluster. **Country** UK. **Source of funding** Voluntary/charity - Mental Health Foundation Government - Department of Health.	**Participants** Children and young people - Sexually abused girls aged 6–14 years. **Sample characteristics:** • Age – Mean age 10 (sd=2.2). • Sex – All participants were female. • Ethnicity – White 45, Black Caribbean 8, Chinese 5, Mediterranean origin 4, unknown origin 2. • Religion/belief – Not reported. • Disability – Not reported. • Long term health condition – Not reported. • Sexual orientation – Not reported. • Socioeconomic position – Not reported. • Type of abuse – Sexual abuse. Total 30 girls were abused by a parent, 28 had more than one abuser, 39 had more than 10 abuse incidents and 27 had experienced abuse of more than 2 years duration. Total 12 girls had experienced touching with or without clothes, 27 genital touching or simulated intercourse, 32 oral, anal or vaginal penetration. • Looked after or adopted status - 49 living with families of origin 19 in	**Effect sizes** **Children and young people's health and wellbeing** There are several difficulties with reporting findings from this paper: - Effect sizes have only been reported where they exceed 0.5 - the reviewing team query the validity of this. - It is unclear whether standard deviations reported in the tables refer to the raw scores, or increase/decreases in scores - Potentially linked to the above, when the reviewing team has attempted to recalculate effect sizes based on data provided in the tables, different estimates are obtained. - Results tables show only changes in scores across measures, rather than raw data. In Table 3 it is states that 'high scores represent the most impairment', yet scores appear to increase for both groups at each time point, suggesting that impairment worsened throughout the intervention. Text elsewhere in the document suggests that this is a typo. A - Univariate analysis 1. K-SADS Data do not appear to be reported 2. Kiddie Global Assessment Scale Note: Effect sizes have been calculated for these variables as none are reported in the paper. Year 1 follow up: Both individual and group therapy conditions showed an improvement in scores, with no	**Overall assessment of internal validity** - **Overall assessment of external validity** + However, lack of reporting in relation to ethical approval is a concern. **Overall validity score** - Lack of clarity regarding statistical analysis, including calculation of effect sizes, and unclear presentation of data are significant limitations in this study.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	foster placements 3 in children's homes • Unaccompanied asylum seeking, refugee or trafficked children – Not reported. **Sample size** Comparison numbers - Individual therapy, n=36 Intervention numbers - Group therapy - n=35. Sample size - n=71. **Intervention** Describe intervention Group therapy plus carer work: Psychoeducational and psychotherapeutic group led by 2 co-therapists. Up to 18 group sessions, each involving around 5 girls. Each group was focused on a pre-arranged topic, with notebooks, task sheets and play materials. Supervision was provided to co-therapists after each session. Carers were worked with separately, sometimes in carers' group but these 'had some difficulties' (p238). Describe comparison intervention Individual therapy plus carer support: Individual weekly therapy sessions lasting 50 minutes, for up to 30 sessions. Early sessions (first phase) were an engagement phase, next 15 sessions focused on issues relevant	significant difference between groups (effect size calculated by reviewing team d=0.07) Year 2 follow up: Both individual and group therapy conditions showed an improvement in scores, with no significant difference between groups (effect size calculated by reviewing team d=-0.05) Exit from study: Both individual and group therapy conditions showed an improvement in scores, with no significant difference between groups (effect size calculated by reviewing team d=-0.02) 3. PTSD Note: We have used effect sizes as reported in the paper for these variables. 3.1 Re-experiencing of traumatic events Year 1 follow up: Both individual and group therapy conditions showed an improvement in scores, with greater improvements in the individual therapy group with medium effect size (d=0.60) Year 2 follow up: Both individual and group therapy conditions showed an improvement in scores, with greater improvements in the individual therapy group with medium to large effect size (d=0.79) Exit from study: Both individual and group therapy conditions showed an improvement in scores, with greater improvements in the individual therapy group with medium effect size (d=0.65). 3.2 Persistent avoidance of stimuli Year 1 follow up: Both individual and group therapy conditions showed an improvement in scores, with greater improvements in the individual therapy group with medium to large effect size (d=0.66) Year 2 follow up: Both individual and group therapy conditions showed an improvement in scores, with	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	to the particular child, final 10 focused on separation and ending. Carers seen by social workers every 2 week. Supervision provided every other week. **Outcomes measured** Children and young people's health and wellbeing outcomes Children assessed at baseline, one year and 2 years. Measures used: 1. Schedule for Affective Disorders and Schizophrenia for School-age Children (K-SADS) (Chambers et al. 1985) 2. Assessment of global impairment of functioning (social, psychological or school functioning) using Kiddie Global Assessment Scale (K-GAS) (Chambers et al. 1985) 3. PTSD scale (Orvaschel 1989).	greater improvements in the individual therapy group with small to medium effect size (d=0.36) Exit from study: Both individual and group therapy conditions showed an improvement in scores, with greater improvements in the individual therapy group with medium effect size (d=0.60). B- Multivariate analysis It is unclear how this has been carried out, particularly as only one F term has been calculated for each dependent variable, despite two independent variables being included in the analysis (condition and placement type). Table 4 appears to present significant data only. 4. PTSD dimensions 4.1 Re-experiencing of traumatic events The relationship between intervention condition and re-experiencing, with K-GAS impairment measure (presumably at baseline) entered as a covariate found that being in the individual therapy group was significantly related to improvements on the re-experiencing trauma scale at first year follow-up ($F=4.3$, $p<0.05$), second year follow-up ($F=7.5$, $p<0.01$) and at exit from the study ($F=7.3$, $p<0.03$). 4.2 Persistent avoidance of stimuli The relationship between intervention condition and persistent avoidance of stimuli, with K-GAS impairment measure (presumably at baseline) entered as a covariate found that being in the individual therapy group was significantly related to improvements on the re-experiencing trauma scale at first year follow-up ($F=5.5$, $p<0.03$).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		C - Change irrespective of therapy The study found that there was a significant improvement in overall functioning in both groups over time (as measured by K-GAS) with a substantial shift from the category of 'major' or 'serious' impairment at baseline to lesser categories in year 1 (chi-square=76.0, p<0.001), and across all time points (chi-square test for most severe 2 categories, and least severe two categories at year one =14.8, p<0.001; chi-square test for most severe two categories, and least severe two categories at year two=29.9, p<0.001). The study also found that there was a significant shift (as measured using McNemar's test) in the numbers of young people with general anxiety (p<0.01, no effect size reported), depression (p<0.001, no effect size reported) and separation anxiety (p<0.001, no effect size reported).	

Research question 17 – What is the impact of social and psychological interventions responding to female genital mutilation? (Prevention of impairment)
No eligible studies found.

Research question 18 – What is the impact of social and psychological interventions responding to forced marriage? (Prevention of impairment)
No eligible studies found.

Research question 19 – What is the impact of social and psychological interventions responding to child trafficking? (Prevention of recurrence, prevention of impairment)
No eligible studies found.

Research question 20 – What aspects of professional practice support and hinder effective response to children and young people who are experiencing, or have experienced, child abuse and neglect?

Research question 20 – Critical appraisal tables

1. **Beckett H, Brodie I, Factor F et al. (2013) 'It's wrong … but you get used to it' - A qualitative study of gang-associated sexual violence towards, and exploitation of, young people in England. London: Office of the Children's Commissioner for England**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. The research team conducts individual interviews with young people (n=150); 11 focus groups with professionals (n=76); and 8 single-sex focus groups (n=38). The comprehensive methods section details the rationale for interviewing participants because of the sensitive nature of the topic and to follow an ethical protocol. In addition, safeguarding concerns have been explored. **Is the study clear in what it seeks to do?** Clear. The forward from Sue Berelowitz, Chief Executive, Office of the Children's Commissioner details the context of the research: very little is known about the prevalence of sexual violence and exploitation within gangs by children and young people against other children and young people. The purpose is to understand through interviews	**Is the context clearly described?** Clear. Under each direct quote, it is clear where data was collected, whether they are a young person or professional and age of participant (if individual interview). The individual interviews with young people (n=150) contain detailed characteristics, however the focus groups held with professionals (n=74) and young people (n=38), it is unclear on the characteristics of these participants. **Was the sampling carried out in an appropriate way?** Appropriate. Very clear that participants were recruited via agencies that were supporting young people to minimise risk. The authors state the potential for 'bias into the sample - and excludes other potential participants with valid contributions to offer - it was felt that the risks of engaging those outside of ser-	**Does the study's research question match the review question?** Yes. The study explores 150 young people's and 76 professional's responses to gang-associated sexual violence and exploitation. The purpose is to understand the prevalence and experiences of young people: chapter 4 is relevant to research question because it explores barriers that hinder young people formally disclosing. **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval was gained from 4 different Research Ethics Committees and relevant local approvals were obtained within each research site. The research team was accountable to a Research Project Advisory Group, a Young People's Advisory Group and local Multi-agency Advisory Groups in each research site.	**Overall assessment of internal validity:** ++ Very comprehensive, effective study with detailed findings that are summarised into recommendations. The methodology has dealt appropriately with the ethics and risk associated with the subject matter, and the research team have ensured the voice of the child is at the heart of the report. **Overall assessment of external validity:** ++ The study meets all criteria and has dealt effectively with ethical considerations. In addition, the interview guide was co-produced with the YPAG to make questions age appropriate. **Overall validity rating:** ++ An excellent, thorough empirical study which meets its research aim and details implications for

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
with young people and professionals' experiences to better inform national and local policy. **How defensible/rigorous is the research design/methodology?** Defensible. Very thorough research design and methodology which was governed and reviewed by a number of different bodies: Research Project Advisory Group; Young People's Advisory Group' and Site specific Multi-agency Advisory Groups. Qualitative interviews were conducted with 150 young people; 11 focus groups with 76 professionals; and eight single-sex focus groups with 38 young people. There is a detailed breakdown of the 150 young people who participated in individual interviews, however the focus group held with professionals and young people is not descriptive. **How well was the data collection carried out?** Appropriately. Data collection section is thorough and the research team explained the measures to ensure the participants comfortability by facilitating the young people to talk in the third person, unless they wanted to actively choose otherwise, i.e. conversational manner using the	vices could not be adequately negated within a time-limited, large-scale, multi-site project such as this' (p12). **Were the methods reliable?** Somewhat reliable. The data is collected by one method, which were qualitative interviews. **Are the data 'rich'?** Rich. The research team cite references to where each finding was collected which helps contextualise responses to each participant. There are limitations as explored: 'Due to the flexibility built into the interviewing process, not all issues were covered with all of these interviewees' (p14). **Is the analysis reliable?** Reliable. Qualitative interviews were thematically analysed using NVivo 8, which underpin the findings in the research. The research team explain the executive decision to generally prioritise the young persons' voice to be presented in the report. **Are the conclusions adequate?** Adequate. The narrative findings of the voice and experience of	**Were service users involved in the study?** Yes. In order to use age-appropriate research questions, the Young People's Advisory Group co-produced the interview schedule. **Is there a clear focus on the guideline topic?** Yes. The relevant section is Chapter 4.2 where young people and professionals state factors that hinder disclosure: confusion about what actually constitutes sexual violence and exploitation; the acceptance of sexual violence and exploitation; and low levels of reporting and seeking support from professionals, i.e. judgement by others, lack of faith in services, perception of police and absence of conviction. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. 150 Young people's experience of gang-associated sexual violence and exploitation, and professionals (n=76) who have experience/specialism working with sexual violence and exploitation.	practice and policy on a local and national level.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
interview schedule as a framework for discussion. There is effective consideration of the commitment to maintaining participants' confidentiality and anonymity. An ethical protocol was developed on the basis of 'no harm should come to any individual as a result of their agreement to facilitate or take part in the work' (p12).	participants contextualise the current knowledge and prevalence of sexual violence and exploitation in gangs. The relevant section to disclosure (Chapter 4) concludes that from the aim of 'identifying learning for embedding more effective systematic response to these issues in the future ... Prompted responses to these are now presented in the form of recommendations' (p51). The recommendations are structured to address national and local policy, which in the context of presenting findings from 6 different localities in England, map the issue with scope to respond.	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Young people were selected because they were/had received support from services, and professionals from statutory services were interviewed i.e. social care, police, and education. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Chapter 4 relates to barriers to professionals for young people disclosing sexual violence and exploitation. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. 'The research aimed to consider: the scale and nature of gang-associated sexual violence and exploitation in 6 areas of England; the main pathways into gang-related sexual violence and exploitation for young people living in these neighbourhoods; and potential models for an effective multi-agency response to the issue' (p6). **Does the study have a UK perspective?** Yes. Study is carried out in 6 different research areas.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		For confidentiality purposes the sites are not named but do 'reflect a broad range of experiences of working with gangs and different demographic profiles' (p6).	

2. **Burgess C, Daniel B, Scott J et al. (2012). Child neglect in 2011: an annual review by Action for Children in partnership with the University of Stirling. Watford: Action for Children**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. The authors conduct 12 focus groups with 114 professionals to across 6 local authorities. The aim of the focus group was to 'gather more in-depth information about prevalence, recognition and response in relation to neglect' (p25). Participants were in-formed in advance as to topics of discussion at focus group. The author states that meetings were recorded and detailed notes were taken. There is no theoretical discussion as to the purpose of conducting focus groups. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. The annual	**Is the context clearly described?** Not sure. Little information. **Was the sampling carried out in an appropriate way?** Not sure. Not reported. **Were the methods reliable?** Reliable. Data is collected by 2 methods - focus group and surveys - and findings are justified within the data collected. **Are the data 'rich'?** Mixed. The annual review has stated that participants in focus group and survey represent a broad range of agencies, however without a breakdown of representatives, it is difficult to distinguish where the information came from. At present, data appears anecdotal and there is no context to the narrative findings.	**Does the study's research question match the review question?** Yes. The relevant section in the report is part two where data from professionals is gathered about what would help neglected children. **Has the study dealt appropriately with any ethical concerns?** No. Not reported. **Were service users involved in the study?** No. Service users did not co-produce this report. **Is there a clear focus on the guideline topic?** Yes. What helps support young people. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes.	**Overall assessment of internal validity:** - The annual review has carried out 12 focus groups which include 114 representatives from different agencies, however the findings and conclusions are 'somewhat convincing' because there is difficulty in identifying or contextualising who said what. There is no consideration of limitations or theory underpinning focus groups. **Overall assessment of external validity:** + Overall, study meets most of the quality criteria however the study is not co-produced and there is no ethical consideration. **Overall validity rating:** - The annual review meets the aim

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
review is descriptive about the local authorities and participants (n=117) who were invited to focus groups, however a limitation is that there is not a clear sample method or demographic information so it is difficult to make generalisations. **How well was the data collection carried out?** Not sure/inadequately reported. Not enough detail on structure of data collection.	**Is the analysis reliable?** Somewhat reliable The project team state that data has been 'analysed in depth to look for emerging themes in the same way as the qualitative information from the survey above' (p26). The detail in the survey analysis is that open-ended questions were grouped under over-arching themes and headings, with particular points of interest highlighted, as it was not possible to include everything. This process of analysis is not underpinned by theory or is not rigorous and the author has not considered the methods to be subject to bias. **Are the findings convincing?** Somewhat convincing. Similar to the analysis, without a clear framework identified and the authors stating that they could not 'include all the detailed information' (p26), it is difficult to form a base of judgement on whether the findings are reliable. **Are the conclusions adequate?** Somewhat adequate. Again, the conclusions highlight overarching themes but there is no consideration of limitations or clarity where	The population of the annual review, in part, is based on professionals working across agencies experience and views of what helps and hinders support services working with neglected children. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Intervention is delivered across all agencies that come into contact with children, i.e. schools and children's social care. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Relates to response. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. The aim is to 'gauge the current situation with regard to neglect and monitor the effects of changes in national and local policy' (p5). **Does the study have a UK perspective?** Yes. The review is carried out across 6 local authorities within the UK.	through the research design and mixed method data collection approach. The findings are representative of a large sample of professionals that work with children who are at the frontline for identifying and responding to child neglect. However, there is little information about consent of participants or what geographical region data is collected, so caution to generalise. Conclusions are difficult to see as reliable because the analysis is 'somewhat reliable'. In addition no ethical consideration. Furthermore, the findings are relatively brief including anecdotal accounts of unspecified respondents, so challenge is contextualising data.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	data is from, i.e. police officer or social worker, hence making conclusions difficult or reliable.		

3. **Children's Commissioner (2015) Protecting children from harm: A critical assessment of child sexual abuse in the family network in England and priorities for action. London: Office of the Children's Commissioner for England**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Qualitative comp 1 Which component? Call for evidence. **Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. **Is the process for analysing qualitative data relevant to address the research question?** Yes. All evidence analysed within overarching research framework. **Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. No contextualisation given to responses to call for evidence, for example which part of the country they came from, or what type of service.	**Quant comp descript A (including incidence or prevalence study without comparison group; case series or case report)** Which component? Survey of adult survivors of intra-familial child sexual abuse. **Is the sampling strategy relevant to address the quantitative research question (quantitative aspect of the mixed-methods question)?** Unclear. Does not state how participants for the survey were recruited. **Is the sample representative of the population under study?** Unclear. There is no analysis of this in the report. **Are measurements appropriate (clear origin, or validity known, or standard instrument)?** N/A.	**a. Does the study's research question match the review question?** Partly. Overall aim of the study is to assess the scale and nature of child sexual abuse in the family environment in England. Some aspects of this are relevant to our review question. We have extracted data in relation to recognition and disclosure (sections 14.1, 14.2 and 14.3) and the impact of intervention (section 15.3), as these relate most closely to our review questions. **Has the study dealt appropriately with any ethical concerns?** Partly. The ethical approach for the study is outlined in Appendix A. However, this did not include getting ethical approval for the study. There is also no mention of how informed consent was obtained from participants. There is some description of how data will be stored.	**Overall assessment of internal validity** - **Overall assessment of external validity** + **Overall validity score** - Little methodological information provided, particularly regarding survey distribution, response rates and representativeness of resulting sample. Limited consideration of ethical issues in reporting.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants? N/A. **Qualitative comp 2** Which component? Site visits. **Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. **Is the process for analysing qualitative data relevant to address the research question?** Yes. Standard pro forma used to collect and analyse data. **Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. Little contextualisation of any differences between sites. **Is appropriate consideration given to how findings relate to**	**Is there an acceptable response rate (60% or above)?** Unclear. No response rate reported - unclear how many individuals were asked to complete the survey. **Is the mixed-methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and quantitative aspects of the mixed-methods question?** Partly. Mixed method design appropriate, but study does not make it clear what the relative contributions of different aspects were expected to be (e.g. oral evidence hearings compared to focus groups). **Is the integration of qualitative and quantitative data (or results) relevant to address the research question?** Yes. **Is appropriate consideration given to the limitations associated with this integration, such as the divergence of qualitative and quantitative data (or results)?**	**Were service users involved in the study?** No. Service users involved as participants, but do not appear to have been involved in designing, conducting or interpreting study. No mention of service users on advisory panel of independent experts. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Partly. Views given by adult survivors and professionals working with sexually abused children and young people. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study includes information relevant to recognition (Q6) and response (Q20). **(For views questions) Are the views and experiences reported**	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
researchers' influence; for example, though their interactions with participants? No. **Qualitative comp 3** Which component? Oral evidence hearings. **Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. **Is the process for analysing qualitative data relevant to address the research question?** Yes. Analysis according to themes identified elsewhere in the research. **Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. No consideration given to, for example, differences in perspective between voluntary and statutory organisations. **Is appropriate consideration given to how findings relate to**	No. Little consideration of limitations of survey approach in general.	**relevant to the guideline?** Partly. Although important to note that, due to age of some of the survivors involved in the research, experiences of services may reflect past service arrangements and practice. **Does the study have a UK perspective?** Yes. England.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
researchers' influence; for example, though their interactions with participants? No.			

4. **Cossar J, Brandon M, Jordan, P (2011) 'Don't make assumptions': children's and young people's views of the child protection system and messages for change. London: Office of the Children's Commissioner**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Defensible. This approach also allowed the children to expand on areas of personal interest to them which they thought important. **How well was the data collection carried out?** Appropriately. Methods of data collection included individual activity-based interviews carried out by adult researchers and one workshop run by a combination of adult and young researchers. A group of young people in one of the participating authorities was consulted by the research team about the design of the recruitment materials.	**Is the context clearly described?** Clear. The United Nations Convention on the Rights of the Child (UNCRC); the UK Children Act 1989; Working Together to Safeguard Children 2010 UK. **Was the sampling carried out in an appropriate way?** Somewhat appropriate. Children and young people fitting the criteria were referred by the 2 participating agencies. A flyer and covering letter for the child was sent in a letter addressed to the parent. Follow-up phone calls were made by local authority workers who passed on the names of children and families who were willing to take part in the research to the research team. The vast majority of participants were recruited at this stage. **Were the methods reliable?** Reliable. Individual activity-based	**Does the study's research question match the review question?** Yes. Children and young people's views of the child protection system. **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval for the research was obtained from the University of East Anglia's School of Social Work and Psychology Ethics Committee. Consent of the parent was sought before the child was approached. **Were service users involved in the study?** Yes. Participants were children under protection plan. A group of young people in one of the participating authorities was consulted by the research team about the design of the recruitment materials. All the young people attending received a £20	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	interviews, workshops. Workshop materials and methods were developed in collaboration with the young researchers who delivered the workshop alongside adult researchers. **Are the data 'rich'?** Rich. **Is the analysis reliable?** Reliable. Interview data and workshop material were recorded and then transcribed, and analysed qualitatively using thematic analysis. In addition to the interview transcript researchers wrote notes about their visit to each child. The qualitative data analysis software package was used (NVivo 8). A coding guide was developed on themes arising from a detailed consideration of 2 of the interviews. Interview data were coded by 1 of 3 researchers, allowing further analysis of key themes and preserving a sense of the complexity of each child's situation. Some basic facts, such as how many children had seen their child protection plan, were gathered from the interviews and these were entered on SPSS. Themes from the workshop were written up and arranged according to the research questions.	voucher in recognition of their time and effort. **Is there a clear focus on the guideline topic?** Yes. Children and young people's views of the child protection system. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children and young people under child protection plan. **Is the study setting the same as at least 1 of the settings covered by the guideline?** No. Not clear (this is not an intervention study). Interviews carried out in homes of the children. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Views and experiences of children under child protection plan. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes. One local authority and one London borough, UK.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Are the findings convincing? Convincing.		

5. **Devaney J (2008) Inter-professional working in child protection with families with long-term and complex needs. Child Abuse Review 17, 242–61**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Defensible. **How well was the data collection carried out?** Appropriately. A semi-structured interview schedule was devised to guide the interview process. The interviews lasted on average 75 minutes (range: 40–105 minutes) and were digitally recorded and subsequently transcribed into a word processing package. The data were then subjected to a content analysis, assisted by the computer software package QSR NVivo 2.0.	**Is the context clearly described?** Unclear. **Was the sampling carried out in an appropriate way?** Not sure. Purposive sampling **Were the methods reliable?** Somewhat reliable **Are the data 'rich'?** Rich. **Is the analysis reliable?** Reliable. Content analysis. **Are the findings convincing?** Convincing. **Are the conclusions adequate?** Adequate.	**Does the study's research question match the review question?** Yes. Exploring views and experiences of professional practice in child protection. **Has the study dealt appropriately with any ethical concerns?** Yes. Approval for the study was provided by the Ethics Committee of Queen's University Belfast in 2002. To protect the identity of interviewees, generic titles have been used in this paper to indicate the role and discipline of the respondent. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. Exploring views and experiences of professional practice in child protection. **Is the study population the same as at least 1 of the groups**	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		covered by the guideline? Yes. Child welfare professionals. **Is the study setting the same as at least 1 of the settings covered by the guideline?** No. Setting not relevant in this context **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Professional practice. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. Views and experiences of child welfare professionals. **Does the study have a UK perspective?** Yes. Belfast, NI.	

6. Franklin A and Doyle L (2013) Still at risk: a review of support for trafficked children. London: The Children's Society

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Qualitative comp 1 Which component? Qualitative in-depth interviews with children and young people identified as trafficked or suspected trafficked.	**Quant comp descript A (including incidence or prevalence study without comparison group; case series or case report)** Which component? Online surveys of local authorities.	**Does the study's research question match the review question?** Partly. To assess mechanisms in place to support trafficked or suspected trafficked children and the role of professionals	**Overall assessment of internal validity** + **Overall assessment of external validity** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question? Yes. In-depth interviews with trafficked children and young people. **Is the process for analysing qualitative data relevant to address the research question?** Yes. Data collected from face-to-face qualitative interviews with children were digitally recorded, fully transcribed and then thematically coded and analysed. Interviews lasted a maximum of an hour, and again with the child's permission, supplementary information about their individual circumstances and contacts with agencies were collected from the voluntary sector staff working with them. Time was spent with the child to ensure that they understood and agreed with any information shared. Each child who took part was given £20 in recognition of the time they had given, and all travel expenses were reimbursed. **Is appropriate consideration given to how findings relate to the context, such as the setting,**	**Is the sampling strategy relevant to address the quantitative research question (quantitative aspect of the mixed-methods question)?** Unclear. Not reported. **Is the sample representative of the population under study?** Unclear. Not reported. **Are measurements appropriate (clear origin, or validity known, or standard instrument)?** Unclear. Not reported. **Is there an acceptable response rate (60% or above)?** No. 34%. **Is the mixed-methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and quantitative aspects of the mixed-methods question?** Yes. Mixed methods: 1. Qualitative in-depth interviews with children and young people identified as trafficked or suspected trafficked 2. Online surveys of local authorities 3. Telephone interviews with key stakeholders.	**Has the study dealt appropriately with any ethical concerns?** Yes. Approved by The Children's Society ethics committee. With the child's permission, supplementary information about their individual circumstances and contacts with agencies were collected from the voluntary sector staff working with them. Time was spent with the child to ensure that they understood and agreed with any information shared. Each child who took part was given £20 in recognition of the time they had given, and all travel expenses were reimbursed. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. To assess mechanisms in place to support trafficked or suspected trafficked children and the role of professionals. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Trafficked children, child protection professionals, e.g. social workers etc.	**Overall validity score** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
in which the data were collected? Yes. To offset the fact that some children might be intimidated, interviews were undertaken in an informal manner, conducted by a worker from within the organisation through which the young person had been recruited. An ethically robust and supportive environment was created to minimise any possible distress to the young person, and they were interviewed at project locations as these are a known safe, supportive and comfortable environment for them. **Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?** Yes. To offset the fact that some children might feel intimidated, the interviews were undertaken in an informal manner, conducted by a worker from within the organisation through which the young person had been recruited. An ethically robust and supportive environment was created to minimise any possible distress to the young person, and they were interviewed at project locations as these are a	**Is the integration of qualitative and quantitative data (or results) relevant to address the research question?** Yes. The integration of quantitative and qualitative data (results) to identify areas for improvement and highlights issue requiring further research. Survey data cleaned and analysed using SPSS (Statistical Package for Social Science). Data collected from the telephone interviews and face-to-face qualitative interviews with children fully transcribed and then thematically coded and analysed. **5.3 Is appropriate consideration given to the limitations associated with this integration, such as the divergence of qualitative and quantitative data (or results)?** Unclear.	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Partly. Children were interviewed at project locations as these are a known safe, supportive and comfortable environment for them. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Child trafficking. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes. Children trafficked into the UK from Burundi, Democratic Republic of Congo, Ethiopia, Ghana, Guinea, Ivory Coast, Nigeria and Vietnam and a South American country.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
known safe, supportive and comfortable environment for them. **Qualitative comp 2** Which component? Telephone interviews with key stakeholders. **Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. Telephone interviews data. **Is the process for analysing qualitative data relevant to address the research question?** Yes. Data collected from the telephone interviews were fully transcribed and then thematically coded and analysed. **Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** Unclear. **Is appropriate consideration given to how findings relate to**			

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
researchers' influence; for example, though their interactions with participants? Unclear.			

7. **Ghaffar W, Manby M, Race T (2012) Exploring the experiences of parents and carers whose children have been subject to child protection plans. British Journal of Social Work 42: 887–905**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Qualitative interviews are appropriate because the researchers aimed to gain an in-depth account of experiences. **Is the study clear in what it seeks to do?** Clear. The aim was to 'explore parents' experiences of child protection systems, including information provided, assessments, case conferences and core groups; their experiences of consultation and support; and their views about factors impacting on their ability to parent their children' (p891). The findings offer implications for practice. **How defensible/rigorous is the research design/methodology?** Defensible. Thorough research design that received detailed scrutiny and approval from the Ethics Committee at one of the universities involved. Confidentiality was	**Is the context clearly described?** Clear. Very comprehensive consideration of demographic context of research participants including details that impacted on their parenting, hence statutory intervention: domestic abuse (n=25); mental health (n=25); drugs/alcohol (n=21); management of child's behaviour (n=18); financial problems (n=17); disability (n=12); childhood adversity (n=6); community (n=5); and housing issues (n=4). **Was the sampling carried out in an appropriate way?** Appropriate. Eligible parents were compiled by each of the 3 local authorities to maintain confidentiality. Reasons cited for not being eligible were ongoing court proceedings, threat of violence or serious health issues. The purposive sampling was to broadly represent the diversity	**Does the study's research question match the review question?** Yes. The study has a direct relevance to responding to child abuse and neglect because it seeks to explore parents' experiences whose children have been subject to child protection plans. **Has the study dealt appropriately with any ethical concerns?** Yes. The research design received detailed scrutiny and approval from the Ethics Committee at 1 of the universities involved in the study. **Were service users involved in the study?** No. Service users did not co-produce the study. **Is there a clear focus on the guideline topic?** Yes. Parents and carers whose children have been subject to child protection plan.	**Overall assessment of internal validity:** ++ Study meets most criteria with clear, balanced findings, analysis and conclusion. Methodology is thorough and the data is collected from large sample group (n=47). **Overall assessment of external validity:** ++ The study meets the topic question. **Overall validity rating:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
upheld by having a contact with each local authority who comprised a list of eligible parents. Each interview followed a semi-structured approach, and parents could choose the location of the interview and all requested to be interviewed in their own home. **How well was the data collection carried out?** Appropriately. The author's state that the study is based on uncorroborated self-report data from parents/carers, however given the research aim to explore the parents/carers perspective on child protection systems this is appropriate.	of local populations and to include both mothers and parents. **Were the methods reliable?** Reliable. Qualitative interviews with 47 participants (from 42 families) where researchers used a detailed interview schedule prepared and piloted prior to the start of the research project. Little detail is provided on what the pilot consists of. **Are the data 'rich'?** Rich. Narrative findings describe parents and caregivers (n=47) whose children have been subject to child protection plans' positive and negative experience of statutory intervention. The data is clearly described in so far as whose perspective is presented i.e. mother/father/couple. **Is the analysis reliable?** Reliable. Research team used N-Vivo software to enable systematic coding of data and reliability was ensured by three researchers examining coding themes emerging from transcripts. **Are the findings convincing?** Convincing. Presentation of find-	**Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Parents and caregivers, child abuse and neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Sample is from 3 local authorities in North England. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Child protection. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. The aim is to conduct qualitative interviews with 42 families (n=47 parents/caregivers) and the response they received from statutory services - 3 local authorities in Northern England. **Does the study have a UK perspective?** Yes - 3 local authorities in Northern England.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	ings are suitably in the child protection process, where data is collected from participants about their experience. Additionally, findings are balanced, ensuring both positive and negative experiences are illuminated in text depending on how many participants agreed/disagreed. **Are the conclusions adequate?** Adequate. Large sample group and corroborate previous studies with similar findings (Baistow and Hetherington 1998). Limitations are recognised in discussion as 'no claims are made that the views of parents interviewed in this study are generally representative of parents in the child protection system' (p901).		

8. Hackett A (2013) The role of the school nurse in child protection. Community Practitioner 86: 26–9

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. The authors conduct 6 individual interviews with school nurses' and school nurse team leaders to determine what is currently known about child protection and aims to identify gaps in knowledge. The study is clear that it was limited by time	**Is the context clearly described?** Clear. Participants are based in a school in 2 geographical areas in Scotland. **Was the sampling carried out in an appropriate way?** Appropriate.	**Does the study's research question match the review question?** Partly. Some findings relate to the review question: nurses discussion of role confusion in child protection as well as skills to help respond to child abuse and neglect	**Overall assessment of internal validity:** + Thorough research design and analysis process that explores school nurses perception on safeguarding duties. The conclusions make recommendations for further

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
and could not interview more professionals' despite having 11 volunteer participants. **Is the study clear in what it seeks to do?** Clear. The study has a clear aim and provides a detailed implications and recommendations for practice. **How defensible/rigorous is the research design/methodology?** Defensible. The methodology is rigorous: detailed description of the research participants (school nurses, n=6); where the participants were chosen and why there is limited interviews (time constraints); ethical approval processes, despite not needing to do so due to the study considered as a service evaluation; how interviews were carried out; and the topic guide adopted to aid thematic analysis. **How well was the data collection carried out?** Appropriately. Interviews were semi-structured and conducted between February and March 2012 at various health board premises which typically lasted between 45 minutes to 90 minutes. The authors provide a	The study includes a detailed methods section that states the research participants were school nurses'/nurse leaders' (n=6), which explains their qualification time and education. In addition, the author's state how they were limited by time so could not use all the 11 volunteer participants and used purposive sampling from 2 geographical regions. **Were the methods reliable?** Somewhat reliable. Data is collected by 1 method - qualitative face-to-face interviews. **Are the data 'rich'?** Mixed. Findings are presented into 3 key themes: role confusion; learning in practice; and moving forward. The paper is 4 pages long so very short, which means findings are limited and selective. The data supports each finding. **Is the analysis reliable?** Reliable. 'The data were analysed thematically. Data analysis commenced at the end of the first interview when notes were made and transcription began. The completed transcription were read several times to aid familiarity and re-	concerns. Moreover, the study focusses on training needs and learning processes. Due to the study being small scale and undertaken in only 1 health board area in Scotland, transferability and generalisability of the findings are limited. **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval was not required as the study was considered to be service evaluation. Permission was granted by the local NHS, and the university research and ethics committee. **Were service users involved in the study?** No. The study was not co-produced. **Is there a clear focus on the guideline topic?** Yes. The study aim is to explore school nurses' perception when identifying and responding to child protection concerns within educational setting. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Professionals: School nurses' and school nurse team leaders.	training needs. There are limitations in that it is a small sample of school nurses (n=6) and open to bias due to the representation of participants from only 2 health boards, not the whole of UK. **Overall assessment of external validity:** + The paper has met most of our criteria, however as a small scale study and brief findings with limited transferability and generalisability due to one locality in Scotland, the conclusions and implication for practice are relevant to area of study. **Overall validity rating:** + Paper is limited and small scale (n=6), however comprehensive research design that links findings with recommendations. Caution to generalise findings.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
detailed topic guide to elicit participant's perspective on current knowledge/training/perception of safeguarding and the school nurses role. Interviews were digitally recorded and transcribed verbatim by researcher. Notes were also made by research team after each interview to explore themes and assist in data analysis process.	duce the data aiding the identification of themes. Upon completion of this stage a coding index was developed and applied to the transcripts. Subsequently, thematic charts were developed for each of the themes. The final stage involved mapping out the key themes and interpreting and making sense of the data' (p27). **Are the findings convincing?** Convincing. The focus of findings is to address additional training needs for the locality where the research is conducted. For the purposes of a 'service evaluation', qualitative interviews are useful to determine gaps in knowledge and examples of best practice. However, as a small sample, generalisability and transferability is limited. **Are the conclusions adequate?** Adequate. The conclusion presents learning recommendations and implications for practice based on data collected from small sample of school nurses (n=6). The limitations are 'the small sample is open to criticism in that it may be biased, as school nurses from one Community Health Partnership area did not	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Education setting. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. The study relates to school nurses' perception of child safeguarding in an educational setting. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. The study details school nurses' perception on what helps and hinders responding to child protection concerns within an educational setting. The findings demonstrate where school nurses experience role confusion and effective ways of working. **Does the study have a UK perspective?** Yes. Scotland.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	participate in the study and their perceptions may have been different' (p29).		

9. Harper Z and Scott S (2005) Meeting the needs of sexually exploited young people in London. London: Barnardo's

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. Unclear how practitioner participants were identified. **How well was the data collection carried out?** Somewhat appropriately. Interview protocol provided for practitioners but not young people. Unclear how interviews were recorded.	**Is the context clearly described?** Clear. **Was the sampling carried out in an appropriate way?** Not sure. Little information given regarding how individual participants were sampled. **Were the methods reliable?** Reliable. **Are the data 'rich'?** Mixed. Relatively good data from practitioner interviews, but analysis of input from young people relatively brief. **Is the analysis reliable?** Not sure/not reported. No mention of double coding of analysis. **Are the findings convincing?** Somewhat convincing. **Are the conclusions adequate?** Somewhat adequate.	**Does the study's research question match the review question?** Partly. Overall research questions do not match our review question, but contains some relevant information and has been included due to overall paucity of evidence on child sexual exploitation. **Has the study dealt appropriately with any ethical concerns?** Partly. Thorough ethical protocol covering informed consent, confidentiality, recording and storing data. However, no ethical approval sought. **Were service users involved in the study?** No. Service users involved as participants, but not in designing, conducting or interpreting study results. **Is there a clear focus on the guideline topic?** Yes.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** + Study was conducted in 2005, which means the findings may be somewhat outdated as awareness of, and practice in relation to, CSE has changed considerably since that time. Relatively sparse reporting of interviews with children and young people.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Is the study population the same as at least 1 of the groups covered by the guideline? Yes. Children and young people at risk of or experiencing child sexual exploitation and professionals working with them. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Contains information relevant to Recognition and Response. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes. England. However, important to note that study was conducted in 2005, since which time there has been much greater awareness of CSE, and significant changes to practice.	

10. Izzidien S (2008) I can't tell people what is happening at home: domestic abuse within South Asian communities - the specific needs of women, children and young people. London: NSPCC

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Seeking views and experiences through interviews and focus groups. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. 30 interviews with professionals, however not sure how many interviewed individually or in a focus group. Sixteen interviews with 2 youth groups, again little information on how interviews were carried out. **How well was the data collection carried out?** Not sure/inadequately reported. Not reported adequately.	**Is the context clearly described?** Not sure. Not reported adequately and difficult to know agencies of professionals as appears to generally be NSPCC domestic abuse services staff. No contextual information on young people and their findings are bullet point summaries. **Was the sampling carried out in an appropriate** way? Not sure. Unsure, not clear how sampling was done. **Were the methods reliable?** Somewhat reliable. One method - interviews with young people through focus group and interviews with professionals through interview and focus group, though not clear. **Are the data 'rich'?** Not sure. Not reported contextually so difficult to determine the narrative and where themes emerged. **Is the analysis reliable?** Not sure/not reported. Not reported.	**Does the study's research question match the review question?** Yes. Views and experiences of South Asian children and young people affected by domestic abuse, as well as practitioners and managers' experience of responding to needs. **Has the study dealt appropriately with any ethical concerns?** No. Not reported. **Were service users involved in the study?** No. Not co-produced. **Is there a clear focus on the guideline topic?** Yes. Children and young people affected by domestic abuse. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children and young people, as well as professionals working in the field of domestic abuse services. **Is the study setting the same as at least 1 of the settings cov-**	**Overall assessment of internal validity:** - Not enough information on methodology. **Overall assessment of external validity:** + Clear relationship with guideline topic and question, however no reports on gaining ethical approval or consent from participants. **Overall validity rating:** - Relevant findings to research question, however with little methodology it is difficult to contextualise the experiences of service users. Additionally, young peoples' voice is lost in the report as only summarised at the end of chapter 3.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Are the findings convincing? Somewhat convincing. Clearly presented narrative findings with supported statements from practitioners and managers. However, it is difficult to contextualise where the information is collected. Furthermore, young peoples' voice is lost in a summary at the end of chapter 3 so no direct quotes. **Are the conclusions adequate?** Adequate. Findings and conclusions link well with implications for further practice and training needs to support young people and children suffering in South Asian communities with domestic abuse. Good insight into the cultural barriers and supported by practitioners experience of supporting service users with these needs.	**ered by the guideline?** Yes. Domestic abuse services in voluntary sector. As well as a domestic abuse youth service. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Child abuse and neglect. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. Response to children and young people affected by domestic abuse and their experiences of seeking help as well as difficulties in responding to this group of service users' needs. **Does the study have a UK perspective?** Yes. England and Wales.	

11. Kazimirski A, Keogh P, Kumari V et al. (2009) Forced Marriage Prevalence and Service Response. London: Natcen

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Qualitative research forms part of a wider mixed methods study, but we have extracted data from qualitative element only.	**Is the context clearly described?** Not sure. An anonymised description of each case study local authority is given in the Methods section. When reporting findings, differences between the local authorities are reported, but not	**Does the study's research question match the review question?** Partly. Part of study is looking at prevalence of forced marriage (not relevant to review question) but part is looking at 'how services are currently responding to cases of	**Overall assessment of internal validity:** + **Overall assessment of external validity:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is the study clear in what it seeks to do? Clear. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. Good justification for selection of local authority case study areas. Less clear how individual participants were sampled and recruited. **How well was the data collection carried out?** Appropriately. Use of topic guide, and all interviews digitally recorded.	linked back to the initial description (e.g. referring to them as local authority A etc.). **Was the sampling carried out in an appropriate way?** Not sure. Little information given regarding sampling approach. **Were the methods reliable?** Somewhat reliable. Only interview data used - not triangulated with other sources of data. **Are the data 'rich'?** Rich. Good exploration of different types of perspectives, although no direct quotes from participants used in research. **Is the analysis reliable?** Reliable. Thematic analysis using specialist software, which allowed checking of extent to which interpretations of the data were shared across the research team. **Are the findings convincing?** Convincing. **Are the conclusions adequate?** Adequate.	forced marriage' (p11). There is content relevant to recognition, early help and response. **Has the study dealt appropriately with any ethical concerns?** No. No ethical approval gained. No description of how consent was gained from professionals involved in the research. Whilst there are less risks involved in interviewing professionals, some may have been directly affected by issues around forced marriage, so consideration of consent and support would have been beneficial. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Professionals working with young people at risk of, or experiencing, forced marriage. However, it should be noted that some of the professionals also worked with adults who were at risk of, or experiencing forced marriage. However, the majority of the report is	Good relevance to question, but no consideration of ethical issues. **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		concerned with practice in relation to children and young people. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least one of the activities covered by the guideline?** Yes. Study has content relevant to recognition, early help and response. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes.	

12. **McGee H, Garavan R, de Barra M et al. (2002) The SAVI report: Sexual Abuse and Violence in Ireland. Dublin: The Liffey Press in association with Dublin Rape Crisis Centre**

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
Objectives of the study clearly stated? Yes. **Design** **Measures for contacting non-responders?** Not reported.	**Describes what was measured, how it was measured and the results?** Yes. **Measurements valid?** Yes.	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval from Royal	**Overall assessment of internal validity** + Lack of disaggregation of childhood versus adulthood abuse presents a challenge in interpretation of results.

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
Research design clearly specified and appropriate? Yes. **Clear description of context?** Yes. **References made to original work if existing tool used?** Yes. Some items were based on previous research, e.g. rape attitudes (e.g. Burt 1980), PTSD measures (Andrykowski 1998). **Reliability and validity of new tool reported?** Not applicable **Survey population and sample frame clearly described?** Yes. Sample frame - whole population of Ireland, sampling by random digit dialling. **Representativeness of sample is described?** Yes. **Subject of study represents full spectrum of population of interest?** Yes.	**Measurements reliable?** Yes. **Measurements reproducible?** Yes. **Basic data adequately described?** Yes. **Results presented clearly, objectively & in enough detail for readers to make personal judgements?** Partly. Results not always clearly disaggregated in terms of those who had experienced abuse as children, and those who had been sexually assaulted in adulthood. **Results internally consistent?** Yes. **Data suitable for analysis?** Yes. **Clear description of data collection methods and analysis?** Yes. **Methods appropriate for the data?** Yes.	College of Surgeons in Ireland. Study monitoring group provided ongoing oversight. Arrangements in place to ensure confidentiality and to manage participant distress, including offer of a follow-up call. **Were service users involved in the study?** No. Service users involved as participants only. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least one of the groups covered by the guideline?** Partly. Survey was of general population, and included adult survivors of childhood sexual abuse. **Is the study setting the same as at least one of the settings covered by the guideline?** Yes. **Does the study relate to at least one of the activities covered by the guideline?** Yes. Data in relation to Response have been extracted here.	**Overall assessment of external validity** ++ **Overall validity score** + Lack of disaggregation of childhood versus adulthood abuse presents a challenge in interpretation of results.

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
Study large enough to achieve its objectives, sample size estimates performed? Yes. **All subjects accounted for?** Yes. **All appropriate outcomes considered?** Yes. **Measures for contacting non-responders?** None reported. **Response rate** 71.44%.	**Statistics correctly performed and interpreted?** No statistical analysis conducted. **Response rate calculation provided?** Yes. **Methods for handling missing data described?** Yes. Some imputation of missing data. **Difference between non-respondants and respondants described?** No. **Results discussed in relation to existing knowledge on subject and study objectives?** Yes. Study also draws on research and other studies undertaken in Ireland. **Limitations of the study stated?** Yes. **Results can be generalised?** Yes. **Appropriate attempts made to establish 'reliability' and 'validity' of analysis?**	**Does the study have a UK perspective?** No. Republic of Ireland.	

Internal validity - approach and sample	Internal validity - performance and analysis	External validity	Overall validity rating
	Not applicable. Not a psychometric survey. **Conclusions justified?** Partly. Some findings cannot be disaggregated in to child and adult abuse when this would be helpful, for example regarding experiences of using services.		

13. McNaughton Nicholls C, Harvey S, Paskell C (2014) Gendered perceptions: what professionals say about the sexual exploitation of boys and young men in the UK. London: Barnardo's

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Study notes that 'qualitative research enables an in-depth exploration of social phenomena and practices, and is particularly suited to exploring emerging and complex issues (Lewis and McNaughton Nicholls, 2014)' (p13). **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. Purposive sampling of interviews to represent a range of regions of England, types of service, seniority, length of service and gender.	**Is the context clearly described?** Unclear. Little information given about location of services from which interviewees were recruited, and little analysis of diversity in responses according to type of service or type of interviewee. **Was the sampling carried out in an appropriate way?** Appropriate. Purposive sampling to obtain a spread of geographical location, professional settings, gender and levels of experience. **Were the methods reliable?** Somewhat reliable. Little detail regarding how online interviews were conducted.	**Does the study's research question match the review question?** Partly. The study has 4 research questions, 1 of which matches our review question which is: to 'suggest ways in which policy and practice may be able to identify and appropriately respond to male victims of CSE, as well as those at risk' (p13). The other three questions are less relevant to this review question and are: identify perpetration and victimisation processes apparent in male-victim CSE cases known to professionals - explore existing service provision for boys and young men at risk of or experiencing CSE - identify future research priorities (p13).	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + UK study but only part of overall research aim was relevant to our review question. **Overall validity rating:** + Only part of overall research aim was relevant to our review question. Study is of reasonable quality, although limited exploration of divergent perspectives across different types of interviewees.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
How well was the data collection carried out? Somewhat appropriately. Little information given regarding data collection. Some participants were 'interviewed online' (p15) - unclear how this was conducted, and no analysis of the impact of this on data.	**Are the data 'rich'?** Mixed. Little consideration of divergence in perspectives along lines of geography, professional background and so on. **Is the analysis reliable?** Somewhat reliable. Analysis conducted using NVivo software. Unclear if procedures such as double-coding of interviews was used. **Are the findings convincing?** Convincing. **Are the conclusions adequate?** Adequate.	**Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval by NatCen's research ethics committee and clarity regarding how data would be presented and stored. **Were service users involved in the study?** No. The report makes occasional references to a young people's workshop. However, this is not described in the methods section. Any findings reported from this strand have therefore not been extracted here. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Professionals working with boys and young men experiencing, or at risk of, sexual exploitation. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by**	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		the guideline? Yes. Relates to recognition (and response). **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes. England.	

14. Pearce J (2011) Working with Trafficked Children and Young People: Complexities in Practice. British Journal of Social Work 41: 1424–40

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. 'Qualitative research methods were used to explore how practitioners understood "trafficking", focusing on how this influenced their identification of and work with trafficked children or young people' (p1427). Large sample group (n=72) with mixed professionals from various agencies. **Is the study clear in what it seeks to do?** Clear. 'The research aimed to explore the differ-	**Is the context clearly described?** Clear. Paper details practitioners agency and how many are representatives. **Was the sampling carried out in an appropriate way?** Appropriate. Research team explicit that purposive sampling was adopted to include practitioners with direct experience working with trafficked children and young people.	**Does the study's research question match the review question?** Yes. The study explores professionals' perspective of factors that help and hinder effective response to trafficked children and young people. **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval granted by the University and NSPCC Ethics Committee. Identity of children	**Overall assessment of internal validity:** ++ Excellent, thorough study that meets aims and objectives. Authors have cited other studies to discuss the current complexities working with trafficked children through interviewing 72 practitioners. **Overall assessment of external validity:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
ent ways that 'trafficking' 'is understood by a range of practitioners from different service agencies; - look at the obstacles that emerge when trying to identify trafficked young people; - chart the process through which a child or young person first gained access to a support agency; and - identify how the practitioner understood the immediate and longer-term needs of the children and young people concerned' (p1427). **How defensible/rigorous is the research design/methodology?** Defensible. Three focus groups held across 3 sites with practitioners from different agency services (n=65) then research team carried out semi-structured interviews with practitioners (n=7) with purposeful sample technique to identify practitioners who currently work with trafficked children. **How well was the data collection carried out?** Appropriately. Topics were initially piloted with experienced practitioners, then the research team conducted qualitative focus group and interviews. The interviews were digitally recorded and transcribed.	**Were the methods reliable?** Somewhat reliable. Data collected one method - qualitative interviews. **Are the data 'rich'?** Rich. Data collected in 3 areas in England to determine practitioner's knowledge and experience with trafficked children. All interviewees are anonymous which enabled honesty from participants - i.e. 'a cynic might say she was trying to get money out of us' (p1431). **Is the analysis reliable?** Reliable. Authors are explicit that transcripts from focus groups were manually thematically analysed by research team. Transcripts from individual interviews were recorded and analysed using NVivo software. These were then triangulated and the final draft report was circulated among the Advisory Group for comment. **Are the findings convincing?** Convincing. Narrative accounts of practitioners experience and knowledge working with trafficked children and young people where examples of good and bad practice are explored. Findings are supported by other studies.	was anonymised to maintain confidentiality. Participants signed consent. **Were service users involved in the study?** No. Research not co-produced. **Is there a clear focus on the guideline topic?** Yes. Professionals identifying and responding to trafficked children and young people. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Professionals in statutory service. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Local Authority. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Child abuse and neglect. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. Professional's knowledge and experience working with trafficked children and young people.	Relevant study to research question with excellent ethical consideration. **Overall validity rating:** ++ Most criteria are met and very comprehensive, in-depth findings.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Are the conclusions adequate? Adequate. Research aims are met because the data enriches the discourse and identifies gaps in knowledge and practice. The authors conclude that there are still difficulties identifying and working with trafficked children and there needs to be a service review of practice so that this group of children and young people are not overlooked.	**Does the study have a UK perspective?** Yes. England.	

15. **Pearce J, Hynes P, Bovarnick S (2009) Breaking the wall of silence: practitioners' responses to trafficked children and young people. London: NSPCC**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Defensible. Good rationale given for why children and young people not directly involved. Substantial sample size. **How well was the data collection carried out?** Appropriately. Focus groups and interviews guided by a specific set of topics	**Is the context clearly described?** Clear. Contextual data provided for the 3 research sites. **Was the sampling carried out in an appropriate way?** Somewhat appropriate. Purposive sampling used for interviews. Practitioner focus groups and case files appear to be convenience sampled, characteristics of both samples well described. **Were the methods reliable?** Reliable.	**Does the study's research question match the review question?** Partly. Study has a range of research questions. The following were judged to be relevant to our review questions: Recognition: 2. Explore the obstacles that might emerge to identifying the numbers of young people trafficked in the three areas. Response: 7. Identify how the professionals feel these needs are best met. 8. Where possible, identify perceptions of how the children/young people feel these needs are best met. We	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** ++ **Overall validity rating:** ++ Thorough data collection, analysis and reporting.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
and digitally recorded. Data recorded from case files using an agreed template.	**Are the data 'rich'?** Rich. Very detailed analysis, drawing out good distinctions between UK children who are trafficked and children trafficked from abroad. **Is the analysis reliable?** Reliable. Focus group data analysed by 2 members of research team. Thematic analysis cross-checked between 2 staff members. Interview data analysed using NVivo. Case files do not appear to have been double coded. **Are the findings convincing?** Convincing. **Are the conclusions adequate?** Adequate.	considered questions 2 to be relevant to our review question on Recognition, and 7 and 8 to be relevant to our review question on Response. **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval for the research project was given by the University of Bedfordshire, School of Applied Social Studies Ethics Committee and by the NSPCC Ethics Committee (p47). Carried out in accordance with ESRC and British Sociological Association guidelines and Barnardo's Research ethics. **Were service users involved in the study?** No. Children and young people were not directly involved - information gathered via analysis of case files. This was due in part to ethical issues associated with involving them in the research. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Trafficked children and practitioners working with them. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. There is material relating to Recognition and Response. **Does the study have a UK perspective?** Yes. England.	

16. **Rees G, Gorin S, Jobe A et al. (2010) Safeguarding young people: Responding to young people 11 to 17 who are maltreated. London: The Children's Society**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Study seeks to explore young people's views and experiences. **Is the study clear in what it seeks to do?** Clear. Clear research question relating to young people's views on seeking and receiving help. The aim is to better understand the experiences of young people to better meet the needs of young people and improve the safeguarding system.	**Is the context clearly described?** Clear. Characteristics of young people is described thoroughly and the author provides contextual information collected at interviews about the maltreatment. There is a balanced ratio of female/male participants involved in the study (10:14) who represent a diverse locality, background and ethnicity. The limitations of conducting interviews with young people is discussed in-depth where the authors give a clear explanation for their	**Does the study's research question match the review question?** Yes. The aim of the study is to explore 'access to, and initial responses of, services for young people with potential maltreatment … to promote protective responses for this target group' (p7). The section relevant to this review question is entitled young peoples' experience of seeking help. **Has the study dealt appropriately with any ethical concerns?**	**Overall assessment of internal validity:** + (When taking in to account additional info from Jobe and Gorin) The study does not have a rigorous methodology or consideration of limitations. Presentation of information is difficult to ascertain where data is collected making conclusions challenging to draw. In addition, there is discrepancy in young people's age as referred in text to both: 11–17; and 11–18.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Implications for practice are discussed in conclusion. **How defensible/rigorous is the research design/methodology?** Defensible. A qualitative design is appropriate given the research question. **How well was the data collection carried out?** Appropriately. The study says that interviews were recorded and transcribed. Study reports that interviews were carried out face to face, that they asked young people about seeking help, being referred to CSC and subsequent responses, and recorded with young person consent.	decision making process, although it is not clear on their sampling techniques. There is consideration of bias, 'our findings may arguably be a partial representation of events as we are unable to present the views or recollections of any of the professionals' young people refer to' (p432). **Was the sampling carried out in an appropriate way?** Somewhat appropriate. Sample appears to be a convenience sample, rather than aiming to be representative of particular categories, however there is diversity of age, gender and ethnic background within the sample. Unclear how young people were selected for interview. **Were the methods reliable?** Somewhat reliable. The data was collected by 1 method which was qualitative interviews. **Are the data 'rich'?** Mixed. The authors state they have explored participants that represent a different locality, background and ethnicity, however there is no recognition of how different young people access a service i.e. the study includes 5 Unaccompanied Asylum Seekers	Yes. The research had ethical approval from the Institute for Research in the Social Sciences Ethics Committee, University of York and the Association of the Directors of Children's Services. **Were service users involved in the study?** No. Service users did not contribute or co-produce the research. **Is there a clear focus on the guideline topic?** Yes. The population of the study is qualitative interviews with young people who discuss their experience of disclosure. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. The participants of the study are 11–17 year old young people who are experiencing statutory support for maltreatment. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Local authority. **Does the study relate to at least 1 of the activities covered by**	**Overall assessment of external validity:** ++ Study relates to question of exploring young peoples' views and experiences of response from children's services and what helped/hindered effective support. **Overall validity rating:** + The study is suitable for scope and the findings enrich discussion about barriers to young people disclosing sexual abuse. Drawing on additional information from Jobe and Gorin (2013), where the research design is more informed, the findings are more convincing as data is richer and analysis is clearer.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	with little recognition that they might have a different experience of disclosing. The narrative findings of the young people make it difficult to distinguish where the information came from. The authors do not state the number of young people that experience what and how so difficult to contextualise and responses are not compared or contrasted across groups. **Is the analysis reliable?** Somewhat reliable. Authors report that they used NVivo and done a thematic analysis. **Are the findings convincing?** Somewhat convincing There is relatively little presentation of the data analysis on which to base a judgement of whether the findings are reliable. There is also little consideration of diversity in views, for example the experiences of children who had sought asylum. **Are the conclusions adequate?** Somewhat adequate The conclusions draw out overarching themes, without considering diverse experiences within the	**the guideline?** Yes. The aim explores 24 young people (11–17) who have been referred to Children's Social Care Services in England and have received statutory support. The paper has a particular focus on the young peoples' experience of disclosing and seeking help for maltreatment. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. The study includes 1 section entitled 'young people's experiences of seeking help'. Interviews have contributed to the findings which are reported to be divided into 4 categories: the difficulties with seeking help; seeking help from peers; seeking help from family members; seeking help from professionals. The other sections of the paper are not relevant to the current review question. **Does the study have a UK perspective?** Yes. Young people are accessed from 6 English local authority areas and represented a range of ethnic backgrounds and ages.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	group, e.g. the children who had sought asylum.		

17. Richardson Foster H, Stanley N, Miller P et al. (2012) Police intervention in domestic violence incidents where children are present: police and children's perspectives. Policing & Society 22: 220–34

Internal validity - approach and sample.	Internal validity - performance and analysis.	External validity.	Overall validity rating.
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear. To assess the views and experiences of children on police interventions in domestic violence; and police and social workers professional practice in their response. **How defensible/rigorous is the research design/methodology?** Defensible. **How well was the data collection carried out?** Appropriately. Interviews and 5 focus groups (Children and young people), interviews (Police)	**Is the context clearly described?** Unclear. Not reported. **Was the sampling carried out in an appropriate way?** Somewhat appropriate. Young people were contacted via a range of organisations providing support for children who had experienced domestic violence. (No further details on sampling and response etc.). No details on how police officers recruited. **Were the methods reliable?** Somewhat reliable. Limited details available. **Are the data 'rich'?** Rich. **Is the analysis reliable?** Somewhat reliable. Thematic analysis of all interview data using NVivo software to assist with data sorting and storage. Qualitative data analysed thematically using	**Does the study's research question match the review question?** Yes. Views and experiences of children on police intervention in domestic violence and police professional practice in response to domestic violence (his study was a mixed methods study examining quantitative file data, prevalence and nature of domestic abuse and qualitative data - only the qualitative data on views and experiences were data extracted to answer the review question). **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical concerns resulted in the 3 groups of participants (young people, adult survivors and perpetrators of domestic violence) being recruited separately so that members of the participant groups had no relationship to one another. Informed consent procedures were adopted for all interviews and ethical approval was	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** +

Internal validity - approach and sample.	Internal validity - performance and analysis.	External validity.	Overall validity rating.
	standard approaches to qualitative analysis -limited details available **Are the findings convincing?** Convincing. **Are the conclusions adequate?** Adequate.	provided by the University of Central Lancashire's Ethics Committee. All participants are anonymised in this paper. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. Part of the study was to assess views and experiences of children on police interventions in domestic violence and police professional practice in response to domestic violence. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children witnessing domestic violence, and police response on domestic violence. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Assumed to be in police stations. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Child abuse and neglect in the context of domestic violence, professional practice of police officers.	

Internal validity - approach and sample.	Internal validity - performance and analysis.	External validity.	Overall validity rating.
		(For views questions) Are the views and experiences reported relevant to the guideline? Yes. **Does the study have a UK perspective?** Yes.	

18. Skinner T and Taylor H (2009) 'Being Shut Out in the Dark'. Feminist Criminology (4)2: 130–50

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Study seeks to explore experiences of young people and in some instances their parent's perspective of the sexual assault disclosure journey. **Is the study clear in what it seeks to do?** Clear. Interviews lasted 1 hour and all participants understood the purpose of the research to gain: 'a more in-depth understanding of the interviewees' perceptions and experiences of the SARC, the criminal justice process as a whole, and their recommendations for improving future practice' (p135). **How defensible/rigorous is the research design/methodology?** Somewhat defensible. The study	**Is the context clearly described?** Clear. The characteristics of participants are clearly detailed and a representation of their demographics are compared against the majority of service users that the SARC support, e.g. p135 titled 'Representativeness of the Respondents' where the SARC's database of referrals contained details of 201 survivors aged 14–16. **Was the sampling carried out in an appropriate way?** Somewhat appropriate. Through coordinating with a local SARC, the authors have preserved confidentiality of the participants', as well as ensuring comfort through enabling the participants' to choose where they would like to be interviewed. The	**Does the study's research question match the review question?** Partly. Study includes views from young people (n=9) and in 6 instances their parents and caregivers perspective on reporting sexual offences to one police service in England. **Has the study dealt appropriately with any ethical concerns?** Yes. There is no legal requirement to gain parental consent before asking individuals under 16 to participate in social research, however the research team adopted a children's rights position which gives the young person the opportunity to state whether they wish to participate in the research. Participants signed a consent form.	**Overall assessment of internal validity:** + The study postulates that there are limitation in the small scale study (n=9) and its generalisability, however the experiences shine a spotlight on hard to reach groups such as young people who have been sexually assaulted. The sample is from a SARC who has supported the survivors which opens the study to potential bias. **Overall assessment of external validity:** + Overall, study meets most of the quality criteria however caution to generalise the UK as the study is based in 1 area where participants are collected from a Sexual Assault Referral Centre and this is

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
provides a breakdown of the participants although the predominance of female interviewees is a possible limitation. The author is clear where the participants were recruited: a Sexual Assault Referral Centre (SARC) in one area in the UK. The SARC was responsible for recruiting participants' who sent letters to survivors of sexual assault in order to maintain confidentially, that had used their service. Eighteen responded however a limitation is the small scale study interviewed 9 survivors. **How well was the data collection carried out?** Appropriately.	SARC has determined the recruitment of participants' because it is deemed to know the survivors best and the study states 'some youth were eliminated from the pool because the young person had not given permission to contact them or the SARC thought they lacked "sufficient age and understanding", were too vulnerable to take part, or had not yet completed counselling' (p134). Consequently, purposive sampling has a limitation of potential bias. **Were the methods reliable?** Somewhat reliable. Qualitative interviews. **Are the data 'rich'?** Rich. Narratives of young survivors are anonymised and their experiences are explored through the pseudonym, i.e. Tammy or Amy. The participant's demographics and experiences are clearly detailed and it is apparent where participants' positive and negative experiences' are portrayed. **Is the analysis reliable?** Not sure/not reported. Data analysis is not reported.	**Were service users involved in the study?** No. However, the study is conducted with a Sexual Assault Referral Centre. **Is there a clear focus on the guideline topic?** Yes. This paper has a focus on young people reporting sexual assault to the police and their experience of the police recognising and responding to the disclosure. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. The study aims to gather the experiences of young people and their parents on the sexual abuse disclosure journey to the police. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Sexual Assault Referral Centre. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. The study covers the response to young people disclosing sexual assault to the police.	written 13 years ago so unsure if SARC still exist. **Overall validity rating:** + Findings are relevant in part to the research question, however as small scale study there is a limitation in generalising conclusions. The authors' discourse suggests that young people have little participation or choice throughout the disclosure journey and are not informed about the process enough so makes recommendations on the findings presented in the study to be more supportive and inclusive. Furthermore, there is no information about the analysis of findings.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Are the findings convincing? Somewhat convincing. The narrative findings are presented through the disclosure journey of the young survivor disclosing sexual assault to the police. The findings are balanced to present positive and negative experiences, however with little information about analysis, it is unclear where the themes derived. **Are the conclusions adequate?** Somewhat adequate. The conclusions support previous adult rape and victims generally (see Hoyle & Zender 2007), however, because of the young person's age, the overarching premise is the survivor feels disempowered and has little choice in what happened to them throughout the disclosure process. The findings support this with data gathered from interviews with the young people. Due to the small-scale, the conclusions are representative of a small populace.	**(For views questions) Are the views and experiences reported relevant to the guideline?** Partly. The young people and their parents comment upon their experience of disclosing sexual assault once the case has been finalised. **Does the study have a UK perspective?** Yes. SARC/Police based in one area in the UK.	

19. Smeaton E (2013) Running from hate to what you think is love: the relationship between running away and sexual exploitation. Ilford: Barnardo's and Paradigm Research

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Qualitative comp 1 Which component? Consultation with young people with experience of running away and CSE while under the age of 16. **Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Yes. **Is the process for analysing qualitative data relevant to address the research question?** Yes. Thematic analysis of each young person's story, then second-stage thematic analysis of common themes. **Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. Little consideration given to variation in the experiences of young people by gender, age, ethnicity and so on. Also unclear how their localities impacted their experience.	**Quant comp descript A (including incidence or prevalence study without comparison group; case series or case report)** Which component? Consultation with professionals working with children and young people experiencing CSE (survey). **Is the sampling strategy relevant to address the quantitative research question (quantitative aspect of the mixed-methods question)?** Partly. Respondents recruited via the NWG Network's newsletter. No consideration of possible gaps that this strategy might leave. All organisations were voluntary sector organisations (although 2 were based within the statutory sector) - it is not clear in the methodology whether this was intentional or simply a product of who responded. **Is the sample representative of the population under study?** Partly. Poor response rate - 28 organisations responded from 500 contacts - suggesting possible	**Does the study's research question match the review question?** Yes. Study explores professional practice in relation to both recognition (Q6) and response (Q20). **Has the study dealt appropriately with any ethical concerns?** Partly. Study was scrutinised and approved by the Barnardo's Research Ethics Committee. One of the principles of the research is given as 'ensuring informed consent and assent' (p12), however no details are given about how consent was obtained. Consideration is given to any needs/issues raised during the research and how these will be addressed. **Were service users involved in the study?** No. Service users involved as participants, but not in conducting or analysing the research. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups**	**Overall assessment of internal validity** - **Overall assessment of external validity** + Could be more detailed description of how addressed ethical issues. **Overall validity score** - Survey of services is entirely of voluntary sector services, and it is unclear whether interviewed professionals represented a wider range of services - the voluntary sector perspective of the research is not highlighted or justified in the research methodology. Little consideration in the findings of how contextual and demographic factors shape participant responses.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants? No. **Qualitative comp 2** Which component? Consultation with professionals working with children experiencing CSE (telephone interviews). **Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Partly. Professionals working with young people experiencing running away and CSE appropriate. However, unclear if all relevant sectors were represented - job titles suggest that they were mainly in the voluntary sector. **Is the process for analysing qualitative data relevant to address the research question?** Yes. Two-stage analysis. **1.3 Is appropriate consideration given to how findings relate to the context, such as the setting,**	bias in favour of most motivated respondents. **Are measurements appropriate (clear origin, or validity known, or standard instrument)?** N/A. Qualitative questions rather than measures. **Is there an acceptable response rate (60% or above)?** No. Response rate appears to be 5.6% (28 responses from 500 contacts), which is poor.	**covered by the guideline?** Yes. Children and young people at risk of or who have experienced sexual exploitation, and professionals working with sexually exploited young people. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to Q6 and Q20. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes. England.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
in which the data were collected? No. Little consideration of how different localities, types of service and so on affected findings. **Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?** No.			

20. **Stalker K, Green Lister P, Lerpiniere J et al. (2010) Child protection and the needs and rights of disabled children and young people: a scoping study: abridged report. Glasgow: University of Strathclyde**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate - 10 interviews were carried out with key informants to seek views on the effectiveness of policy and practice in meeting the needs of disabled children. The intention was to interview respondents from the whole of the UK, but the author states that there were high profile child abuse cases at the time, so resources were stretched. Eight representatives from Scotland and 2 from England were the final sample. **Is the study clear in what it seeks to do?** Clear. Clear aims and objectives.	**Is the context clearly described?** Not sure. Not reported. **Was the sampling carried out in an appropriate way?** Appropriate. Authors state that the participants were purposive selected to enable the appropriate professionals and policy makers to enrich the data to meet research aims. **Were the methods reliable?** Not sure. Not reported. **Are the data 'rich'?** Not sure. Not reported.	**Does the study's research question match the review question?** Yes. The aim of the interviews were to 'explore current debates and issues regarding child protection and disabled children, seek views on the effectiveness of policy and practice in meeting the needs of disabled children and thus help identify key questions for further research' (p7). **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval was obtained from the University of Strathclyde Ethics Committee. Ad-	**Overall assessment of internal validity:** - No information on data analysis, collection or how findings are contextualised. Does provide an overview of what is currently known and meets research aims. **Overall assessment of external validity:** + Meets criteria. The participants generally represent Scotland (n=8) so caution to generalise the UK. **Overall validity rating:** -

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
How defensible/rigorous is the research design/methodology? Somewhat defensible. Little information on the research design. **How well was the data collection carried out?** Not sure/inadequately reported. Not reported.	**Is the analysis reliable?** Not sure/not reported. Not reported. **Are the findings convincing?** Somewhat convincing. The study seeks to illuminate current debates and issues regarding child protection and disabled people, however as small sample of practitioners working in central government and policy makers (n=10), findings are limited to good and bad practice locally. Additionally, '... there is a high level of consistency in several of the key findings arising from different parts of the study - the research review, the policy analysis and the key informant interviews' (p21). **Are the conclusions adequate?** Somewhat adequate. Little information on how data is collected, analysed and concluded so unsure where themes arise and if it is anecdotal to local practice. The study includes a small section about 'pilot methods of seeking disabled children's views', however this is not reported because it has methodological rather than substantive findings.	ditionally, participants gave consent and choice on whether to partake. **Were service users involved in the study?** No. Service users did not co-produce the report. **Is there a clear focus on the guideline topic?** Yes. Child protection and disability. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Professionals and policy makers working within the child protection field. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Child protection. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Child protection. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. The 10 interviews are relevant because the aim is to explore child	Poor research design with little information about how data is collected, analysed and thus, conclusions are difficult to draw.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		protection for children with disabilities and highlights gaps within the field. **Does the study have a UK perspective?** Yes. Interviews with 2 informants from England and 8 in Scotland. Caution to generalise the whole of the UK as the research team were unable to gain interviews with representatives from Wales or Ireland.	

21. Stanley N, Miller P, Richardson Foster H (2012) Engaging with children's and parents' perspectives on domestic violence. Child and Family Social Work 17: 192–201

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Study seeks to understand service users' experiences of responses to domestic violence. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. Sampling strategy for participants is unclear. No information given regarding demographic characteristics of sample and to what extent this is typical/representative.	**Is the context clearly described?** Unclear. Little information regarding to the 2 research sites. Also, important to note that some young people were recruited from groups outside the research sites - these are also not described. **Was the sampling carried out in an appropriate way?** Not sure. Very little information given regarding sampling. **Were the methods reliable?** Somewhat reliable. Unclear why focus groups used for gathering	**Does the study's research question match the review question?** Partly. Focus of the research is on the process of domestic violence notifications, which relates more to service configuration rather than specific interventions. However, we have extracted data in relation to young people and their caregivers' views on various services. Our data extraction relates only to the qualitative interviews with young people, survivors and perpetrators. **Has the study dealt appropriately with any ethical concerns?**	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity rating:** - Very little information given regarding participants' demographic characteristics. Little description or justification of sampling. Unclear why focus groups used for some participants but interviews used for others.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
How well was the data collection carried out? Not sure/inadequately reported. Little information given regarding questions posed in focus groups or interviews. Unclear why focus groups used for young people and interviews for survivors and perpetrators.	data from young people, and interviews for perpetrators and survivors. **Are the data 'rich'?** Rich. Divergences in opinion are drawn out, and use of illustrative quotes. Authors also report how many people had experience of each service. **Is the analysis reliable?** Somewhat reliable. **Are the findings convincing?** Somewhat convincing. Not enough information provided about interviewees and focus group participants to be able to contextualised responses. **Are the conclusions adequate?** Somewhat adequate. Brief summary section given at the end of the chapter, but does not cover all key issues.	Yes. Ethical approval given by University of Central Lancashire's Research Ethics Committee and the NSPCC's Research Ethics Committee. Procedures for informed consent are described. Support was available to young people and survivors as required. **Were service users involved in the study?** No. Service users involved as participants, but not in design, conduct or interpretation of results. **Is there a clear focus on the guideline topic?** Yes. Study relates to domestic violence in families with children. This relates to guideline as Adoption and Children Act extended definition of significant harm to include 'impairment suffered from seeing or hearing the ill-treatment of another'. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
		Does the study relate to at least 1 of the activities covered by the guideline? Yes. Study relates to response to child abuse and neglect (with child abuse defined as witnessing domestic violence). **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes. England.	

22. Taylor J, Stalker K, Fry D et al. (2014) Disabled children and child protection in Scotland: investigation into the relationship between professional practice, child protection and disability. Edinburgh: Scottish Government Social Research

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Mixed. Study's research questions are clear, but findings are not written up according to the original research questions. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. Participants had to have experience of at least 2 cases of child protection concerns in relation to a disabled	**Is the context clearly described?** Unclear. Little consideration of how respondents' roles and experience may affect their responses. **Was the sampling carried out in an appropriate way?** Not sure. Insufficient information to judge. **Were the methods reliable?** Not sure. Insufficient information regarding focus groups. **Are the data 'rich'?** Mixed. Use of case studies adds to richness of	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Partly. Ethical approval by University of Edinburgh Moray House School of Education Ethics Committee. Mentions informed consent for interview participants, but not focus group participants. **Were service users involved in the study?** No.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Limited information on ethical considerations in relation to focus groups. **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
child, which would seem to be relatively low. The spread of professional backgrounds and extent of experience of recruited participants is not reported. **How well was the data collection carried out?** Somewhat appropriately. Use of Critical Incident Technique as part of in depth interviews. All interviews digitally recorded with consent. Less clear how focus groups were conducted.	data, but overall there is little consideration of reasons for divergences in opinion between interviewees. **Is the analysis reliable?** Somewhat reliable. Results analysed in various ways, including use of modelling. With regard to thematic analysis of qualitative data it is not clear how themes were formulated and checked, and some of the themes identified do not appear to be coherent concepts. **Are the findings convincing?** Somewhat convincing. Themes developed in thematic analysis do not appear to be coherent and distinct. **Are the conclusions adequate?** Somewhat adequate.	**Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Practitioners working with disabled children at risk of or experiencing significant harm. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to recognition and response. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes. Scotland.	

23. Wirtz L (2009) Hidden children: separated children at risk. London: Children's Society

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Interviews with trafficked migrant children, professionals and analysis of 34 case studies.	**Is the context clearly described?** Clear.	**Does the study's research question match the review question?** Yes. Trafficked children, 'hidden' children (i.e. separated children	**Overall assessment of internal validity:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is the study clear in what it seeks to do? Clear. **How defensible/rigorous is the research design/methodology?** Defensible. **How well was the data collection carried out?** Appropriately. Trafficked children Semi-structured interviews, with 4 of the 8 young people interviewed 1:1, and half in a focus group that took place in a pre-existing peer group run by a voluntary agency. Interviews in English without an interpreter. If interviewees had wanted an interpreter, the young person would have been given a choice of a man or woman and they would have had an opportunity to reject the interpreter offered. As English was not their first language, the interviewer used simple language, rephrased questions and repeated and rephrased some answers to check that they were understood. Much of the interview involved drawing and writing if the young people were willing to do so. Four of the 8 young people consented to be audio recorded during interviews, and the interviewer took written notes. Professionals: 13 of	**Was the sampling carried out in an appropriate way?** Appropriate. Interviewees and steering group members were contacted after being recommended by their current or former support worker, some young people have been known to The Children's Society and the rest were referred by other voluntary agencies or schools. **Were the methods reliable?** Reliable. **Are the data 'rich'?** Rich. Based on data in Section 6.3 (Disclosure) and 6.4 (Support needs). **Is the analysis reliable?** Somewhat reliable. Interviews were transcribed and sorted by theme by the interviewer. Written notes were taken by the researcher during steering group meetings (no further details). **Are the findings convincing?** Convincing. **Are the conclusions adequate?** Adequate.	who are exploited by the people responsible for them). **Has the study dealt appropriately with any ethical concerns?** Yes. The research followed The Children's Society's Research Unit's ethical procedures which include the use of a written set of ethical principles and procedures that must be approved by 2 senior researchers from the Research Unit and 2 external academics. The ethical procedures were characterised by the following principles: opt in, informed consent, anonymity, incentives, follow-up, confidentiality, safety and avoiding emotional manipulation. Potential interviewees were told that whether they were involved or not, it would not affect the support they get from the worker who referred them to the project. Also overseen by a Professionals Steering Group, and a Young People's Steering Group of 4 young adults aged in their early 20s who had been hidden children themselves. Their role was to advise on the methodology, ethical procedures, analysis and recommendations. **Were service users involved in the study?** No.	**Overall assessment of external validity:** ++ **Overall validity rating:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
the 15 professionals consented to be audio recorded during interviews, and the interviewer took written notes in the other cases. One professional answered the questions by email. Interviews transcribed and sorted by theme by the interviewer. Written notes were taken by the researcher during steering group meetings. They were asked for case studies, their experiences of good practice when supporting hidden children and their recommendations for improved direct and joint work. Case studies: drawn from a few London boroughs, particularly in East London, and North West England.		**Is there a clear focus on the guideline topic?** Yes. Views and experiences of trafficked migrant children. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Trafficked migrant children and professionals from the voluntary and statutory sectors. **Is the study setting the same as at least 1 of the settings covered by the guideline?** No. Not clear as not reported. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Child trafficking. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. Views and experiences of trafficked migrant children and professionals from the voluntary and statutory sectors. **Does the study have a UK perspective?** Yes. Migrant children trafficked into the UK.	

Research question 20 – Findings tables

1. **Beckett H, Brodie I, Factor F et al. (2013) 'It's wrong … but you get used to it' - A qualitative study of gang-associated sexual violence towards, and exploitation of, young people in England. London: Office of the Children's Commissioner for England**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: 'The research aimed to consider: the scale and nature of gang-associated sexual violence and exploitation in six areas of England; the main pathways into gang-related sexual violence and exploitation for young people living in these neighbourhoods; and potential models for an effective multi-agency response to the issue' (p6). **Methodology:** Qualitative study. The research team adopted a qualitative approach to conduct: - Individual interviews with 150 young people - 11 focus groups with 76 professionals - 8 single sex focus groups with 38 young people. The study took place between 2011 and 2013.	**Participants** Children and young people. Individual interviews - Young people aged 13–28 (n=150) Focus groups - Young people (n=38) Professionals/practitioners - 11 focus groups were conducted with 76 professionals across 6 research sites. Representation from fields of social care, education, health, policing and the justice system, specifically working within the gangs and sexual exploitation/sexual violence. **Sample characteristics** • Age - Interviews - Participants ranged from 13 to 28: Under the age of 18 (49%); 18–20 (28%); 21–25 (21%); and 25–28 (2%). Focus groups - Not reported. Professionals - Not reported. • Sex - Interviews - 52% were male, with 48% female. Focus groups - Not reported. Professionals - Not reported. • Ethnicity - Interviews - The self-reported ethnicity of interviewees: 32% Black/Black British; 28% White; 21% Dual heritage; and 18% Asian/Asian	**Narrative findings** Young people and professionals commented on the current service response, the general sense is the system is still in its early stages. Professionals identified the following challenges when responding to gang associated sexual violence: - One professional said, 'I think that's the crux across the board - the funders that goes into prevention - mentors, youth workers, support workers - is pathetic compared to what goes on in terms of dealing with the consequences' (p46). - In gang affected areas, there is a history of community-based and statutory-based initiatives integration that impact on how to tackle gang violence effectively. - Professionals lack knowledge and awareness of the issue. - Silo working where agencies are practicing separately and not strategising together to work effectively which consequently affects information sharing and partnership working. - Funding issues and short term initiatives. - A historical view point and initiatives to tackle male-male violence rather than cross-fertilising practice to help tackle sexual violence within gangs (see Firmin 2013). - Inadequate data collection/monitoring (p46).	**Overall assessment of external validity** ++ The study meets all criteria and has dealt effectively with ethical considerations. In addition, the interview guide was co-produced with the YPAG to make questions age appropriate. **Overall assessment of internal validity** ++ Very comprehensive, effective study with detailed findings that are summarised into recommendations. The methodology has dealt appropriately with the ethics and risk associated with the subject matter, and the research team have ensured the voice of the child is at the heart of the report.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Country: UK, England. 'To maintain confidentiality and protect participants, the identity of the research sites is not being revealed' (p6). **Source of funding:** Government - Inquiry of the Office of the Children's Commissioner into child sexual exploitation in gangs and groups. Led by the University of Bedfordshire.	British. Focus groups - Not reported. Professionals. • Religion/belief - Interview - Not reported. Focus groups - Not reported. Professionals - Not reported. • Disability - Interview - Not reported. Focus groups - Not reported. Professionals - Not reported. • Long term health condition - Interview - Not reported. Focus groups - Not reported. Professionals - Not reported. • Sexual orientation - Interview - Not reported. Focus groups - Not reported. Professionals - Not reported. • Socioeconomic position - Interviews - Most participants reported that they were in some form of education (45%), training (20%) or employment (18%), with only one in eight identifying as Not in Education, Employment or Training (NEET). Focus groups - Not reported. Professionals - Not reported. • Type of abuse - 87% (n=131) had direct, often multiple connections with gangs. of the 131 participants, 59% were/had been directly involved in a gang (M=70% vs. F=47%); 32% had been gang-associated (M=25% vs. F=39%); 35% had friends/and or family involved; 23% were having/had previously had a 'romantic relationship; with a gang-involved person (all female bar one); 57% had personal		**Overall score** ++ An excellent, thorough empirical study which meets its research aim and details implications for practice and policy on a local and national level.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	experiences of sex and/or relationships in gangs. The remaining 13% (n=19) participants grew up in gang-affected neighbourhoods. Focus groups - Not reported. Professionals - Not reported. • Looked after or adopted status - 38% of participants reported current or previous involvement with children services, although it is not clear what support this was. Focus groups - Not reported. Professionals - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Interviews - Not reported. Focus groups - Not reported. Professionals - Not reported. **Sample size** Interviews - 150 participants. Focus groups - 8 single sex with 38 young people. Professionals - total of 11 focus groups held with 76 professionals.		

2. **Burgess C, Daniel B, Scott J et al. (2012). Child neglect in 2011: an annual review by Action for Children in partnership with the University of Stirling. Watford: Action for Children**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The annual review pro-cess by Action for Children	**Participants** Professionals/practitioners –	**Narrative findings**	**Overall assessment of internal validity:** -

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
in partnership with the University of Stirling seeks to 'gauge the cur-rent situation with regard to neglect and monitor the effects of changes in national and local policy' (p5). The project team collated evidence through a variety of methods: collated statistics from across the UK to record incidence of neglect by bodies such as WHO and UNICEF; analysis of policy developments across 4 nations in the UK; telephone survey administered to 35 local authorities and email survey to 12 local authorities to understand which children come to the attention of professionals and what support/resources are in places; in-depth, on-site focus groups across 6 areas in the UK; online polls undertaken by 2062 adults in the general public to understand views and	Qualitative: n=114 professionals Quantitative: n=47 local authorities **Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** Qualitative sample size: Research team held 12 focus groups with 6 (two areas were combined) local authorities: Local authority 1: n=12 Local authority 2: n=14 Local authority 3+4: n=47 Local authority 5: n=21 Local authority 6: n=20 Total participants = n=114 professionals	Findings are 'Professionals - are children recognised but not helped?': Factors that hinder effective response: - Participants commented that there was a tendency for children and families to 'bounce in and out of services' (p14). - Some felt that when a family had a lot of services involved, focus can be lost and this can impact the child. Additionally, long-term real change is not always apparent. - If families are in rural areas are considered hard to reach. One focus group responded commented: 'In a large rural area transport is an issue for families trying to get to services. It is too expensive for people to get buses, if they exist, and many families can be very isolated. There is already less money allocated for this - we have to fight for it' (p14). - Specialist services for some groups of young people, i.e. children with disabilities transitioning to adult services, and 'in some areas black and minority ethnic groups are not well catered for' (p14). - Resources and re-organisations are impacting on the delivery of regular service, as one respondent accounts: 'Constant service restructuring means staff changes, which are detrimental to both families, who have to make new relationships, and to other agencies who have to relearn who they need to contact with concerns about children. Staff turnover in some services can be a major problem' (p15).	The annual review has carried out 12 focus groups which include 114 representatives from different agencies, however the findings and conclusions are 'somewhat convincing' because there is difficulty in identifying or contextualising who said what. There is no consideration of limitations or theory underpinning focus groups. **Overall assessment of external validity:** + Overall, study meets most of the quality criteria however the study is not co-produced and there is no ethical consideration. **Overall validity rating:** - The annual review meets the aim through the research design and mixed method

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
experiences of awareness of child neglect; and 2174 professionals responded to an online poll to ascertain views of the nature and qualities of their responses and barriers to neglected children. **Methodology:** Mixed methods. The qualitative aspect of the annual review was conducting 12 focus groups across 6 local authorities with a total of 114 participants across a range of agencies: children's services; housing; health service staff; the police; education; and third sector agencies. The focus group data is most relevant to research question as provides in-depth data in response to 'how good are we at recognising children who are at risk of, or are experiencing, neglect?'	Quantitative sample size: 47 local authorities.		data collection approach. The findings are representative of a large sample of professionals that work with children who are at the frontline for identifying and responding to child neglect. However, there is little information about consent of participants or what geographical region data is collected, so caution to generalise. Conclusions are difficult to see as reliable because the analysis is 'somewhat reliable'. In addition no ethical consideration. Furthermore, the findings are relatively brief including anecdotal accounts of unspecified respondents, so challenge is contextualising data.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
The quantitative aspect administered telephone interviews (n=35) and via email in a further 12 to local authorities (t=47). This was a 63.5% response rate as the survey was sent to 74 local authorities in total. The survey asked for information about 'definitions, referral pathways, inter-agency working, initial assessment tools, statistics on those formally identified as neglected and for statistics on "proxy" data, such as that relating to children affected by parental substance misuse, mental health problems or domestic abuse' (p25). **Country:** UK. **Source of funding:** Not reported.			

3. **Children's Commissioner (2015) Protecting children from harm: A critical assessment of child sexual abuse in the family network in England and priorities for action. London: Office of the Children's Commissioner for England**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: Aim to assess the scale and nature of child sexual abuse in the family environment in England. We have extracted data in relation to Recognition (sections 14.1, 14.2 and 14.3) and the Impact of intervention (section 15.3), as these relate most closely to our review questions. **Methodology:** Mixed methods. Study comprised: 1. A call for evidence to collect examples of good practice 2. A DfE dataset request for data on victims and perpetrators 3. Police force dataset request for data on victims and perpetrators 4. Site visits and focus groups in 6 sites, including consultation with 32 agencies and focus groups with 5 victim/survivor organi-	**Participants:** Adult survivors of child abuse - A survey of 756 survivors of child sexual abuse, all were over the age of 18. Professionals/practitioners - Site visits and focus groups in 6 sites, involving 32 agencies Oral evidence hearings with 9 professionals from statutory bodies and 10 professionals from voluntary and community organisations **Sample characteristics:** • Age - Survivor survey: 18–24 n=50, 25–34 n=133, 35–44 n=214, 45–54 n=251, 55–64 n=88, 65+ n=20. Other evidence strands: Age of participants not reported. • Sex - Survivor survey: Female n=483, Male n=51, Unknown n=215, Other n=5, Prefer not to say n=2. Other evidence strands: Sex of participants not reported. • Ethnicity - Not reported. • Religion/belief - Survivor survey: No religion n=283, Unknown n=215, Christian (all denominations) n=196, Other n=42, Jewish n=10, Buddhist n=7, Muslim n=2, Hindu presume n=0 (not shown on pie chart) Other evidence strands: Religion of participants not reported.	**Narrative findings** Data have been extracted from Section 15.3. 15.3 The impact of intervention by statutory and non-statutory services The study reports that respondents to the survivor survey and participants in focus groups reported 'feelings of disappointment and distress' regarding their contact with statutory services. Some reported that they had not been believed. One participant said: 'I had a few sessions of counselling via my GP, this was awful, limited to a couple of sessions and actually ... left me feeling let down yet again. It took me many years to search for a local charity who were absolutely amazing, without them I most probably would not be here today' (p78). (Note: this quote is attributed to an 'interview' - unclear if this refers to focus group? Interviews not mentioned in methodology.) The study reports that a number of evidence strands found that in some cases, victims of intrafamilial sexual abuse are not allowed to discuss the abuse with other members of their family to avoid prejudicing the outcome of criminal justice processes. Where survivors' abuse had been recognised, and services intervened appropriately, the study reports that this had made 'a significant difference' to people (p79).	**Overall assessment of internal validity** - **Overall assessment of external validity** + **Overall validity score** - Little methodological information provided, particularly regarding survey distribution, response rates and representativeness of resulting sample. Limited consideration of ethical issues in reporting.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
sations. 5. Oral evidence hearings with 9 professionals from statutory bodies and 10 professionals from voluntary and community organisations 6. A survey of 756 survivors of child sexual abuse 7. Data request from 4 helplines 8. A rapid evidence assessment of research evidence on intra-familial sexual abuse 9. There is ongoing research with children and young people (assume this is not reported here). The data extracted here are drawn from strands 1, 4, 5 and 6. **Country:** UK, England. **Source of funding:** Government. Office of the Children's Commissioner.	• Disability - Survivor survey: No disability n=397, Disability n=106, Unknown n=211, Don't know n=41, Prefer not to say n=1. Not reported for other strands. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** Survivor survey: 756 Site visits: 32 agencies (unclear how many individuals) Focus groups: 5 focus groups victim/survivor organisations and 3 focus groups with survivors of child abuse (unclear how many Oral evidence hearings: 9 professionals from statutory bodies and 10 professionals from voluntary and community organisations Total sample size unclear.	Total 50% of respondents to the survivor survey reported that their experience of sexual abuse had affected their ability to access health services. 20% of respondents reported that they still avoid going to the GP, and 17% reported avoiding going to the dentist. One respondent wrote: 'Went to the doctors a lot as a child with sore throats hoping they would see something was wrong' (Female survivor, aged 45–54).	

4. **Cossar J, Brandon M, Jordan, P (2011) 'Don't make assumptions': children's and young people's views of the child protection system and messages for change. London: Office of the Children's Commissioner**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To seek children and young people's views of the child protection system and to consider how those views might contribute to improving responses to abuse and neglect. **Methodology:** Qualitative study. Activity-based interviews and workshops. **Country:** UK, London. **Source of funding:** Government - Office of the Children's Commissioner.	**Participants** Children and young people – n=26 **Sample characteristics** • Age – 6–17 years • Sex - 13 girls and 13 boys from 18 families • Ethnicity – Three-quarters of the children were white British. Those from minority ethnic groups included Asian/Asian British, black British Caribbean and black British African children, as well as 2 children who were of mixed heritage. • Religion/belief - Not reported • Disability - 3 of the 26 children had a learning disability or learning needs (mild or moderate learning difficulties or a statement of special educational needs). • Long term health condition- Not reported. • Sexual orientation - Not relevant. • Socioeconomic position - All children living at home with at least one parent. • Type of abuse - Not reported, but all had a child protection plan at the time of interview. • Looked after or adopted status - Not reported.	**Narrative findings** Views and experiences of children and young people of the child protection system: A. What helps the child feel safe? 1. Coping strategies - People to confide in about worries a. friends and family, such as family members and neighbours. b. Professionals, such as social workers, teachers, including a multi-systemic therapist, a pastoral support worker, a substance misuse worker, a counsellor, a psychiatrist and a youth worker; also the police. Younger children (under 12 years of age likely to confide in social workers). 2. Cognitive and behavioural strategies to manage worries a. by compartmentalising his life and keeping things in separate spheres (p39); or did not like to talk about problems because … did not really trust anyone (p39). b. Use harmful strategies to manage their worries, such as anger and violence, depression and self-harm, as a response to their challenging circumstances. Also substance misuse such as drinking heavily. In summary, most of the children had people they could confide in about their worries or from whom they sought help. The most common source of support was friends or family, a range of professionals	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** n=26 children and young people **Outcomes measured** Satisfaction with services Views and experiences of children and young people with the child protection system.	and nearly all the children could identify a professional who had helped them. As a response to challenging circumstances, some children resorted to drinking or self-harming as a way of managing other worries. Implications for professionals in Child Protection: 1. Be aware of the strategies that the child has developed to deal with their worries and the problems in the family (p41). 2. Consider who might be a trusted adult for the child and how they might continue to be involved in their support (p41). B. What is the child's view of the professional concerns about their family? 1. Minimal awareness- some children unaware of professional concerns (p42). 2. Disagreement with concerns - Some children felt that professional concerns were misplaced or no longer relevant because of children's misunderstanding of the nature of the concerns (p42). 3. Partial agreement with professional concerns - viewed their intervention as intrusive, or 'social workers misread the situation' (p43). 4. Disagreement with professional view of parenting - some children felt social workers had misinterpreted relationships within their families and they rejected what they perceived to be the social work view of their mothers. More common for some children to acknowledge concerns that focused on themselves, than to acknowledge concerns relating to parenting (p43).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		5. Professionals underestimate the risk - Some children disagreed with the professional concerns because they felt that professionals were not seeing important aspects of their circumstances; or that child acknowledged that the difficulties in his family were recognised by professionals but felt that he was being held responsible (p44). In summary, children and young people varied in their awareness of the professional concerns. A minority of the children and young people thought that professional concerns were mistaken or unfounded and these tended to be younger children. Some young people agreed that there had been a reason for professionals to be involved with their families but felt that the concerns were now in the past. There was a tendency for the children and young people to disagree particularly with professionals' views of their parents. They were more likely to acknowledge problems with their own behaviours. Two young people thought that there was cause for concern in their families which professionals overlooked (p45). Implications for professionals in Child Protection (pp45–6) 1. Maintain an openness to the child's view of the situation. 2. Where there is a difference between the child's and the social worker's views, make sure that the child's views are represented and the social worker's position is explained to the child. 3. Understand the importance of the child's relationship with their social worker. 4. Make sure that the child is seen on his or her own.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		5. In cases of particularly difficult dynamics between professionals and parents, managers should consider providing a separate worker for the child. C. What is the child's understanding of the child protection system?(pp48–50) The degree of understanding was age-related- 1. Minimal awareness: the majority of children under 12 years of age not understanding the meaning of child protection meeting/plan/conference. 2. Partial awareness – piecing together the jigsaw: might have detailed knowledge of an aspect of the child protection system or know that their parents were in court, but be unsure why, mostly in children up to 13 years of age. 3. Clear understanding: able to give a clear account of the child protection process, mostly older children over 13 years of age who had attended a child protection meeting. In summary, children's understanding of child protection was age-related and rated into 3 categories, minimal, partial and clear understanding, with most of those having a clear understanding being in the older age group. The majority of the children were categorised as having a partial understanding. Children with a partial understanding of child protection sometimes had a detailed account of part of the process. They had some overview of the system but could not give a coherent account. They often relied on parents and siblings for information. Some of the children whose families were involved in court proceedings had a better understanding of the court process than they did of other aspects of child protection. Children with a	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		clear understanding were older and all of them had attended a child protection meeting (p51). Implications for professionals in Child Protection (pp51–2) The vast majority of the children and young people had some understanding of the child protection system, even those under the age of 10. How much information it is appropriate for a young child to have about the formal child protection system is a difficult judgement for professionals and parents to make. 1. Be aware that the child has a view about the child protection process as well as about the problems within the family. 2. Think about the sense that the child makes of the social work intervention and check what they find helpful and unhelpful. D. How much does the child participate in the child protection process? (pp53–5) It depends on: 1. Children's relationships with their social worker: majority of young people knew how to get in touch with their social worker. Older children (aged 12 and over) more likely than younger children to be seen on their own. 2. Positive and trusting relationship with social worker: Some of the children had trusting relationships, based on honesty and trust, with their current social worker and felt that the social worker was working with them; and social workers had succeeded in overcoming the young person's previous negative experiences of workers. 3. Minimal relationship: some children did not have a relationship with their social worker, who was seen as	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		a remote figure who got in contact with their parents occasionally. These children did not see the social worker on their own. 4. Factors affecting the relationship between children and social workers: a. Child felt pressured - Some children viewed their social worker as someone who came round and asked them questions, not as someone to whom they could talk about their worries. They also felt the social worker tried to take over his mum's role. Some children felt that the social worker dwelt too much on problems and negatives and did not see the good things in their family (p55). b. Twisting our words - What put young people off speaking to their social workers was that the social worker misrepresented what they said; or that professionals had exaggerated the situation. c. Confidentiality- Children expressed concern about the information when they spoke to their social workers who did listen but then told everyone what was said. Children felt it important that information was shared appropriately and not 'blabbed' to others who did not need to know (p56). 5. The child protection system (pp57–63) a. Reports and assessments - Some children knew about the reports but did not have a chance to see them, or correct them, as some children saw all or part of reports and assessment and some said that they did not. Few young people talked about discussing reports with social workers before meetings. c. Meetings - Not all children were invited to go to a meeting; some did not want to. Many children were not aware of the choices open to them to have their views expressed at a meeting.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		d. Attending meetings - some children did not feel able to participate by asking questions or being listened to, and in general were dissatisfied with their level of participation, despite that some recalled having spoken at the meeting, and said that they were supported by a family member, a friend, or by a teacher or advocate. e. Preparation - some children had seen leaflets providing information before the meeting or spoken to their social workers and found it helpful; others did not feel well prepared. Some felt unable to say that they disagree with the report during the meeting. f. Support at the meeting - Most children were supported by a family member, or professionals and advocates. g. Speaking at meetings - children felt able to speak at the meeting, encouraged by the chairperson. A minority of children felt they were only partly listened to, suggesting that children's voices were marginalised. Some of the young people found it difficult to be honest at the meeting, particularly when asked 'awkward' questions in front of their parents. h. Decision-making- few young people felt they were 'a lone voice opposing the plan', and not involved in making them. i. The emotional impact of meetings - 'nerve-wracking', 'lairy' when listening to third party (a teacher in this case) giving opinions about the child when they had never previously met. Attendance at a core group was difficult because '... they were just all talking and I didn't understand what they were saying'. j. Feedback- Of those children who were aware of meetings, whether or not they attended, some said the outcomes were explained to them but some said it had not been explained.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		k. Family group meetings- opinions varied, some thought it was largely redundant but some found it productive. In summary, older children more likely to be seen by social workers on their own. Some children had trusting relationships with their social workers and some children reported having minimal relationship with social workers, seeing them rarely or only at meetings. Some children found it difficult to talk to their social workers because they felt pressured by the social worker asking questions, or said that the social worker twisted what they said. Few children saw reports or assessments and it was rare for the young person to have a chance to discuss the report with the social worker. A small minority of children were aware of different ways their views could be given to the meeting. Most of the children who attended the meetings found them difficult because they were being asked awkward questions in front of their parents. Few felt they were listened to and spoke about decision-making at the meeting. Not many children had seen their child protection plans (p63). Implications for practice (pp64–6) For practitioners 1. Ensure children are given information about the child protection process that is appropriate to their needs. In assessing this, and their involvement, take account of the dynamics within the family as well as his or her age and understanding. 2. Ensure that the child has an appropriately worded copy of the child protection plan which should be discussed with the child and incorporates their input.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Consider how best to explain the plan to a young child. For managers 1. Be mindful of the existing guidance on involving children and young people in the child protection process and think about how best to involve each individual child. Include the child in these discussions. 2. Local authorities should recognise the importance of the child's relationship with the social worker and how this contributes to the engagement in the process of help. They should organise the work so that social workers can get to know children, and are not viewed as remote but powerful figures. 3. Promote guidance on good practice so that workers think about how best to involve each individual child. For policy makers Promote guidance on good practice and make it easily accessible to child protection professionals (pp64–6). E. What is the child's experience of intervention? (pp67–74) 1. Child protection investigation - experience with the Police varied, some children found the experience frightening and felt compelled to take part; some found it 'quite supportive'. The sensitivity of the professionals (police or social workers) involved made a difference. 2. The benefits of having a social worker - majority of children found their involvement with the social workers helpful, that the social workers had helped to improve things in their families, e.g., social workers giving practical support (cash support, vouchers for day-outing) and advocating for the children, liaising with	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		other agencies, sharing assessment reports with the children. This was based on establishing trusting and positive relationship between the child and the social worker. 3. Unhelpful aspects of having social worker Many of the children could recall both something helpful and something unhelpful and it was an exception to be either completely damning of social work involvement (p70). 4. Intrusion Some children found it unhelpful when their social workers visited them at school, they felt the child protection process 'controlling' and they felt 'criticised and personally monitored', as 'everything I do when I walk out the house gets reported back to social services'. They felt there were 'rules imposed on them' and they were not involved in making the rules. 5. Increased tension in the family Children aware that the child protection process was stressful for their parents. Having social workers had increased the pressures in the family, such as when undergoing multi-systemic therapy, an intensive family intervention. The professional attempts to change/improve one child's behaviour resulted in increasing the risk to another sibling, making the situation worse, and things deteriorated in the family. This was likely to reflect a negative (minimal and adversarial) relationship between the child, family and the social workers. F. Changing views - Children said their understanding of social work intervention changed over time, sometimes because they had grown older, and agreed that they and the family needed support. G. Stigma - Some children aware of a stigma attached to being involved, and careful whom they told	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		about having a social worker or about having a child protection plan. They did not want to draw attention to themselves, being thought of being 'weird' (pp73–4). In summary, the sensitivity of the professionals involved made a difference to how difficult the experience was for the young people; many children could identify something helpful that their social worker had done for them, such as practical help, improvements in their family relationships, liaison with schools and talking through their problems. Identified advantages of having a child protection plan included extra help at school or getting priority for services. Negative aspects of having social work involvement included intrusion, increased stress within the family, and having to deal with stigma (p74). Implications for practice (p75) 1. To form relationships with children social workers need to be knowledgeable about child development and the impact of abuse and maltreatment. 2. Social workers need to have good skills in communicating with children, based in this knowledge. This should be an important focus of social work training and continuing professional development (p75).	

5. **Devaney J (2008) Inter-professional working in child protection with families with long-term and complex needs. Child Abuse Review 17: 242–61**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To explore the views of experienced child welfare	**Participants**	**Narrative findings**	**Overall assessment of internal validity:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
professionals about families known to the child protection system with long-term and complex needs in relation to the process of intervention with families. **Methodology:** Qualitative study. In-depth semi-structured interviews. **Country:** UK. **Source of funding:** Government - funded by the Research and Development Office of the Northern Ireland Health and Personal Social Services.	Experienced child welfare professionals such as: Social worker (n=2); Social work manager (n=12); Health visitor (n=3); Nurse manager (n=4); Medical profession (n=3); Education profession (n=2); Voluntary sector (n=1); and Police officer (n=1) **Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Participants were child welfare professionals. • Looked after or adopted status - Participants were child welfare professionals. • Unaccompanied asylum seeking, refugee or trafficked children - Participants were child welfare professionals. **Sample size** n=28 child welfare professional, mean length of experience was 14 years (range: three to 26 years)	Views and experiences of experienced child welfare professionals A. Process of intervention by staff 1. Engagement with families and children a. The way staff approached and treated families as being of prime importance in attempting to work effectively with complex cases, can be difficult as often there was a disagreement about the purpose and role of social workers' involvement with families: 'A lot of it is when some parents dig their heels in and say "No, we're not doing that …". I think then you find that you lock horns with families' (social worker) (p250). b. Skills needed to address parents' anxieties and defensiveness: 'I think it probably still boils down to the relationship between family and worker … if they see that they have been listened to, their views have been respected, their views have been heard … they are more likely to go along with that' (family centre manager) (p250). c. Engaging parents could be complicated by staff from different disciplinary backgrounds holding differing views. This could be helped by parents experiencing a seamless tailored service whereby the professionals involved worked closely together: 'I think if the parents see that professionals who are involved in the care of their child are all working together heading towards the same goal, and are making decisions you know are in the best interest of the child, and if they see that there's good communication	**Overall assessment of external validity:** ++ **Overall validity rating:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		between the professional I think it can only help with gaining their cooperation' (paediatrician) (p250). d. Staff from different disciplinary backgrounds holding differing views about how people should be involved in the process: 'I feel … social workers are very open with their clients. Some of the other professionals might be wary about how it's going to affect their role with the family …' (nurse manager) (p251). 'Health visiting, GP and schools … are (now) much more open and honest with us and with clients. That does very much inform decision-making' (social worker) (p251). 2. Multi-disciplinary relationships a. Need for better working relationships between agencies, to promote inter-agency coordination and collaboration in safeguarding children and promoting their welfare: 'Trying to get other professionals on board … is a real stumbling block … there is this view out there that it's social services problem, it's not ours' (social work manager) (p251). b. Important to establish an open and regular communication that moved beyond information sharing to an analysis of the issues in the family and sharing of ideas about how to move forward 'There are people who have lots of information and don't want to share it' (health visitor) (p252). 4. Assessment and decision making: A more in-depth picture of the relationships within the family and any	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		health and social needs of the mother and children in the family was what was needed. a. Quality of information in case notes and records varied from 'so bland that it provided little insight into the family situation' (p253), to 'A tendency to provide a descriptive account of the events rather than a considered analysis' (p253). b. How much non-social services staff were involved in the information sharing and decision making in case conferences 'I think they tend to share information and leave it on our doorstep to deal with … I've never seen a GP or a school being given a recommendation to follow through on the child protection plan. I don't know whether that's perhaps our own culture … the social worker is left to deal with that …' (social worker) (p254). c. Rather than disinterest in the safety or wellbeing of the child or a feeling of wanting to be distanced from the decision making, non-social services staff need to build up their confidence '… some of the new practitioners in health visiting wouldn't be as confident in decision making, because of their lack of skills and their lack of fully understanding the significance of significant harm and … the register, but hopefully … they to come to me for induction … and hopefully through my training and supervision that they'll develop those skills' (nurse manager) (p254). d. Quality of chairing: central role of the chairperson in both setting the tone for the meeting, managing the	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		business and involving participants including family members: 'I think if you've got a good Chair, all things will be taken on board … most of ours would be very succinct, very focused because they know they're only going to have professionals there for a given amount of time, very organised' (paediatrician) (p254). 'I mean there's still a lot of variance in even how decisions are made, how case conferences are run … like there's some you go to and it's so slick and it's so competent and the Chairperson is so appropriate in moving things along and sort of dealing with people's issues and whatever, and other times you come out and you go, that was a complete and utter mess and it was disorganised' (health visitor) (p255).	

6. Franklin A and Doyle L (2013) Still at risk: a review of support for trafficked children. London: The Children's Society

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: 1. To assess the experiences of children identified as trafficked or suspected trafficked and accommodated in local authority care, their understanding about the types of services they had received and how professionals supported them, their understanding of care processes, and transition at aged 18 (face-to-	**Participants:** Children and young people. Trafficked children (definition of trafficked children is children who have been trafficked, and are now outside their country of origin). Professionals/practitioners - Social workers, independent reviewing officers (IROs), directors of Children's Services and other professionals providing care to trafficked children. Also key stakeholders, social care managers and front line social workers, solicitors (welfare and immigration) and representatives from the voluntary sector.	**Narrative findings** A. Recognition The complexity of disclosure and identification of trafficked children (pp24–31) 1. Trafficked children became very confused and frightened following discovery or escape. They might be kept locked up or threatened or controlled which prevented children from escaping, as did threats made against their family. 2. Children may disclose unintentionally, or may wait until they feel safe, or until they have a trusting relationship, or they may reach a point of desperation 3. Trafficked children could have little opportunity to escape their traffickers and exploiters as these chil-	**Overall assessment of internal validity** + **Overall assessment of external validity** ++ **Overall validity score** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
face interviews) 2. To assess mechanisms in place to support trafficked or suspected trafficked children and the role of professionals (telephone interviews) 3. To assess the multi-agency response in the context of best practice in child protection and safeguarding (online surveys) 4. To identify good practice and areas for improvements. **Methodology:** Mixed methods. Face-to-face interviews, survey and telephone interviews. **Country:** UK. **Source of funding:** Government - The Home Office.	**Sample characteristics:** • Age - Trafficked children: aged between 15 and 23 years Professionals Not reported. • Sex – Trafficked children: 15 girls and 2 boys Professionals: not reported. • Ethnicity - Trafficked children: from 9 different countries of origin: Burundi, Democratic Republic of Congo, Ethiopia, Ghana, Guinea, Ivory Coast, Nigeria and Vietnam and a South American country. Professionals: Not reported. • Religion/belief - Trafficked children: Not reported Professionals: Not reported. • Disability - Trafficked children: Not reported Professionals: Not reported. • Long term health condition - Trafficked children: Not reported Professionals: Not reported. • Sexual orientation - Trafficked children: Not reported Professionals: Not reported. • Socioeconomic position - Trafficked children: Not reported Professionals: Not reported. • Type of abuse - Trafficked children: Reasons for being trafficked- domestic servitude (n=7), forced labour and criminal activity including cannabis cultivation and selling drugs (n=3)	dren lacked a clear understanding of what is happening to them (i.e. they have been trafficked), or knowledge of their rights and sources of support available once discovered (p25). 4. Not speaking English and possibly not even knowing which country they are in is also a major barrier. 5. Some trafficked children were criminalised for activities such as documentation offences and criminal acts which they were forced to engage in while being exploited; some were treated as adults when discovered and were subsequently wrongly placed within the adult criminal justice system or immigration detention facilities. 6. Private fostering arrangements might be a concern as these can be used to hide trafficked children (p31). A lack of awareness, understanding and training can lead to some practitioners and the police not identifying trafficked children even in situations where children have sought help. The emphasis of recognition would have to be on adults having awareness of the indicators of trafficking to enable discovery and identification, especially those whose decisions may impact on their care arrangements, such as the police, immigration officers and legal representatives (p31). B. Response (pp37–8) 1. The response to trafficked children when they go missing - There was a general lack of awareness of trafficking meant some children were not properly protected, supervised, accommodated and supported, and went missing (p38). a. Professionals emphasised the importance of multi-agency working to react to trafficked children going missing: quick action had to be taken to minimise risk,	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	and sexual exploitation (n=9). Professionals: Not relevant. • Looked after or adopted status - Trafficked children: Not reported Professionals: Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Trafficked children (see Types of abuse). **Sample size** Trafficked children: n=17 (15 girls and 2 boys) Professionals: On-line surveys (n=30 heads of Local Safeguarding Children's Boards). Telephone interviews (n=18); 9 social care managers and front line social workers, 2 solicitors (welfare and immigration) and 7 voluntary sector staff.	such as developing a multi-agency safety plan, record keeping, securing safe accommodation, working with specialist, trained and supported foster carers, providing intensive one-to-one support and the forming of a trusting relationship with an independent adult. b. Some respondents felt that strategies should include the option to access support beyond the local area, such as the use of reciprocal arrangements between local authorities or a national specialist foster care programme to locate the whereabouts of missing trafficked children (p38). c. A national approach, i.e. establishing a national database of missing trafficked children would help to alert different agencies. d. When trafficked children do go missing, there is a strong possibility that children return to their traffickers and their cases should be considered as abduction cases and treated accordingly by local authorities and the police. e. Training of specialist foster carers to prevent children being placed in inappropriate placements (p38). 2. Use of current guidance and multi-agency working (pp39–47) a. Child trafficking toolkits and NRM guidance on trafficking were considered helpful but some felt there was little understanding of how those indicators should be incorporated in assessment processes, to predict risk and as a way of determining the most appropriate services for a child (p39). b. Good social care for trafficked children should focus on a duty to protect these children, rather than focusing on them as being trafficked. c. Effective multi-agency working was highlighted as being important in providing the right type of support	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		to trafficked children, but multi-agency working was highly dependent on a shared understanding and proper training across agencies, and on the importance placed on the issue by local authorities. Few local authorities had developed multi-agency strategic or operational groups focusing on trafficking and few local authorities had implemented current guidance by developing multi-agency strategic or operational groups focussing on trafficking, or undertaken local needs assessments. d. Key areas identified as working well at a local level included: provision of comprehensive training across faiths and cultures; effective protocols, effective working with local police; establishment of a sub-group of the LSCB; multi-agency strategy meetings about individual cases; appointing specialist workers or commissioning local voluntary sector organisations to provide training and support; developing a communications strategy about private fostering across LSCB partners (p44). e. Some respondents expressed frustration with the NRM (National Referral Mechanism) process and did not see it providing support to trafficked children. f. Trafficked children's experience and distress of having to repeat their story multiple times to multiple agencies indicates that improved multiagency working could be of significant benefit to them (p47). 3. Support for trafficked children (pp47–9) Support given by social services a. Although some individual social workers were seen as supportive, practice varied widely and only a minority of the sample of trafficked children were happy with the care and support provided by their social workers. Although some individual social workers.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		b. Lack of continuity of care, trafficked children often had multiple social workers and had to frequently repeat their story. c. Trafficked children's main criticism of social care support: lack of contact and support, not being listened to and social workers not doing things that they should do, leading to a lack of trust. d. Trafficked children reported seeking the services and support they needed from welfare solicitors and/or support workers from voluntary organisations. e. Stakeholders repeatedly highlighted the need to see what has happened to the child as a child protection issue and not an immigration issue, and to respond accordingly. f. There were concerns that social work teams specialising in one area (e.g. asylum or looked after children) might not have the full range of knowledge or skills required to manage the often complex situations. g. Child protection support could be compromised by some trafficked children's uncertain immigration status especially during transition from children's services to adult services/independence (p49). Age assessments (p66) a. Many trafficked children undergo multiple age assessments, which some practitioners thought were highly problematic for this group of children. b. Age assessments were often taking place in police stations and in some cases they were being undertaken by social workers who were making pre-judgements (p66). c. Children reported they were often not believed during age assessments and the questioning of them	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		made it difficult to have good relationships with their social worker. d. Some children had their age wrongly identified and had been sent to adult prisons, detention centres or been placed in adult accommodation, placing them in a very vulnerable position (p66). Access immigration advice and to specialist services (p56) a. Stakeholders highlighted that access to good quality immigration advice was a concern. b. Local authorities reported barriers to supporting trafficked children: insufficient accommodation, a lack of understanding amongst social workers of the immigration and legal systems and pressures relating to the immigration process. There were also barriers to providing an allocated permanent social worker to trafficked children. Access to generic services: local authorities faced some difficulties in accessing appropriate education, mental health services and leisure opportunities for trafficked children (p62). Therapeutic support: Some trafficked children had experienced serious mental health issues as a result of the exploitation they had been exposed to and reported that they had benefited from counselling accessed through their social or key worker, solicitor or voluntary sector workers (p59). Access to education a. Education for trafficked children was seen as vitally important, although provision was varied. Local au-	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		thorities faced some difficulties in accessing appropriate education, mental health services and leisure opportunities for trafficked children (p62). b. Some trafficked children received incorrect advice about their education, and/or did not receive their right to an education (p66).	

7. Ghaffar W, Manby M, Race T (2012) Exploring the experiences of parents and carers whose children have been subject to child protection plans. British Journal of Social Work 42: 887–905

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The aim of the study was to explore 42 families in 3 local authorities in North England whose children have been subject to child protection plans. **Methodology:** Qualitative study – 42 semi-structured, qualitative interviews were conducted with parents. **Country:** UK, 3 local authorities in Northern England. **Source of funding:** Not reported.	**Participants** **Sample characteristics** • Age - Most parents/carers interviewed were in their twenties (n=20); thirties (n=12) or forties (n=7). Two were under 20 and one was over 50. The mean age for mothers when their first child was born was under 20 for 2 authorities and 21.8 in the third authority. • Sex - Female - 39 Male - 8 • Ethnicity - White British = 36 British; Asian = 5; and Duel heritage = 1. • Religion/belief - Not reported. • Disability - Disability was not the focus of the study, but it is reported that a third of participants were affected by disability. Three parents reported having a disability themselves, and 11 reported having a child with a disability. • Long term health condition - Not reported.	**Narrative findings** Findings are presented in the child protection process so themes can be categorised as: 1. Information provided 2. Experiences of assessment 3. Case conferences and Child Protection Plans 4. Consultation and decision making 5. The role of the professional. 1. Information provided: - Families reported that they were not routinely given written information about child protection procedures which impacted on their ability to compete on equal terms. - Some parents commented on not understanding the information provided. As one parent recalls, 'It was all in double Dutch. I attempted to read it, but it didn't make sense, it was like reading a doctor's prescription' (p897). - In some instances, parents did not recognise the seriousness or purpose of the child protection processes. One parent commented: 'they didn't really make it clear what was actually happening' (p897). 2. Experiences of assessment:	**Overall assessment of internal validity:** ++ Study meets most criteria with clear, balanced findings, analysis and conclusion. Methodology is thorough and the data is collected from large sample group (n=47). **Overall assessment of external validity:** ++ The study meets the topic question. **Overall validity rating:** ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Sexual orientation - Not reported. • Socioeconomic position - Not reported beyond 'out of the six ethic minority parents, none of them were employed' (p893). • Type of abuse - On the child protection plans, children the category of registration: - sexual abuse= 6 - neglect = 23 - physical abuse = 9 - emotional abuse = 6 (2 joint). • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** 47 parents and caregivers from 42 families.	- Many parents commented that they felt social workers did not acknowledge the level of stress experienced during the assessment process. - Families (n=4) stated they concealed information from professionals for fear of consequences, i.e. domestic abuse, mental health issues or drug taking might impact on the removal of their children. - 10 parents felt the deficit model of assessment was disempowering, as one parent recalls the assessment report submitted to a case conference, 'There was nothing positive, it was all bad. When you're in a room full of professionals it's not very nice'. - Conversely, one mother accounts her strengths being recognised and this empowered her and improved morale e.g. 'They told me ... I've got potential to do it. I've just got to get my mind in the right place' (p898). 3. Case conferences and Child Protection Plans: - 37 parents commented upon their daunting experience of case conference, emphasised by feeling unable to present their perspective, e.g. 1 parent said that case conferences were 'very heavy and quite draining. I used to feel ill when I came out' (p898). - On the other hand, several parents mentioned the positive experience of the conference chair who was supportive. - The study asked if parents agreed (n=19) or disagreed (n=17) with their Child Protection Plan decision. Reasons cited for agreeing were that parents felt able to access more services or in domestic abuse instances, safety. However, parents who disagreed felt that they did not fully understand the safeguarding responsibilities of professionals. 4. Consultation and decision making:	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		- 18 parents were positive about being included in the consultation process, whereas 6 parents felt limited in decision making but still felt listened to. One parent felt, 'I wasn't involved in any decisions, but they explained (things) very well, and they listened' (p899). 5. The role of the professionals - these findings are representative of what parents felt help and hindered effective response from professionals. Factors that helped: - Parents (n=32) considered positive experiences when social workers had good listening skills, were open about agency involvement with clarity. - Parents recognised the supportive and practical function of social services, as a couple who disagreed with agency involvement remarked, 'they had (baby)'s best interest at heart ... they did the job properly' (p900). - It was considered effective if social workers spent time with the children. - Additional social worker qualities included good organisation and reliability. Factors that hindered: - 19 families highlighted that the change of social worker made them feel uncomfortable having to divulge personal information to a new worker. - 12 families experienced professionals who lacked empathy. - There were some examples of parents feeling stigmatised by social workers because of their substance misuse.	

8. Hackett A (2013) The role of the school nurse in child protection. Community Practitioner 86: 26–9

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The aim of the study is to explore school nurses' perceptions of their role in child protection; identify skills required to undertake this role; identify training needs. **Methodology:** Qualitative study. Qualitative interviews with 6 school nurses in a Scottish city. Purposive sampling to comprise 3 school nurse team leaders and 3 school nurse staff from 2 geographical areas. Interviews were semi-structured, using an interview topic guide that explored the participants 'current knowledge, experience, qualification and training; perceived role and responsibilities relating to safeguarding and protection children and young people; current involvement in child protection issues;	**Participants** Professionals/practitioners - Individual interviews were conducted with School nurse team leaders (n=3) and school nurse staff (n=3) from 2 different geographical areas. Length of qualification ranged from 5–37 years, while experience in school setting ranged between 3–19 years. Most participants were degree educated (n=5), with one completing a nurse diploma level and 2 participants held the specialist practitioner qualification. **Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size**	**Narrative findings** Findings have been thematically arranged into 3 key themes: 1. role confusion; 2. learning into practice; and 3. moving forward. 1. Role confusion: - Participants commented on a lack of clarity over the school nurses responsibility, albeit all knew the need to follow child protection concerns. One participant described being unsure of her role, 'I said that I was quite clear what my role is, safeguarding and disease prevention, and then sometimes I act like I am unsure about what my role is, you know what I mean, because I am pushed somewhere else' (p27). - All participants wanted clarity to roles and responsibilities in order to manage expectations of families, other professionals and managers. - There were varying accounts of a school nurses role in child protection. One participant felt that '… it depends how far you want to be involved with child protection … when a child protection case conference comes up that is my devotion to that case and to see it through'. Whereas, one participant commented '… it is not social concerns. We are very much I think physical health … we don't do home visits' (p27). 2. Learning into practice - This is not relevant as it addresses training needs which will not be covered in the review question. 3. Moving forward	**Overall assessment of internal validity:** + Thorough research design and analysis process that explores school nurses perception on safeguarding duties. The conclusions make recommendations for further training needs. There are limitations in that it is a small sample of school nurses (n=6) and open to bias due to the representation of participants from only two health boards, not the whole of UK. **Overall assessment of external validity:** + The paper has met most of our criteria, however as a small scale study and brief findings with limited transferability and generalisability due to one locality in Scotland, the conclusions and implication for practice are

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
perceived education/training needs for identifying and responding to child protection issues; awareness of child protection courses/training; and perceived/actual barriers to uptake of training/educational courses' (p27). **Country:** UK, Scotland.	Total participants = 6.	- All participants commented that confidence and communication skill were paramount in taking forward child protection concerns. - Two participants commented that listening was important too, '… hearing and listening are probably slightly different, but it is being open to hear what they are saying … take it step by step' (p28).	relevant to area of study. **Overall validity rating:** + Paper is limited and small scale (n=6), however comprehensive research design that links findings with recommendations. Caution to generalise findings.

9. Harper Z, Scott S (2005) Meeting the needs of sexually exploited young people in London. London: Barnardo's

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The aim of the study was to understand: 1. The nature and extent of sexual exploitation in London 2. The service needs of young people at risk of sexual exploitation 3. Gaps in existing service provision in London 4. Examples of promising practice which could be shared across London. The	**Participants:** Children and young people - 12 young people aged between 13 and 19. Professionals/practitioners - Interviews with a range of practitioners including child protection co-ordinator (n=32), police (n=10), health service (n=10), education service (n=2), local authority looked-after children's service (n=2), residential home manager (n=1), youth offending team (n=3), secure unit manager (n=5), specialist sexual exploitation service (n=6), voluntary sector service with	**Narrative findings** 1. Police - Barriers to an effective police response Barriers identified included: - No clear lead on child sexual exploitation - Lack of clarify regarding remit between Child Protection Units, Community Safety Units, clubs and vice, Sapphire teams, local borough police, CID, Missing Persons Units, Public Protection Police and the Child Abuse Prevention Unit. - Lack of resources for this area of work. - Unable to prioritise without sufficient resources, or if there is no intelligence to suggest a problem in relation to sexual exploitation.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** + Study was conducted in 2005, which means the findings may be somewhat outdated as

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
study also includes examples of facilitators and barriers to identification of, and response to, child sexual exploitation, which is what our data extraction has focused on. **Methodology:** Qualitative study. Qualitative interviews with young people and practitioners as part of a wider study which also included audit. **Country:** UK, England. **Source of funding:** Voluntary/charity - Corporation of London's Bridge House Trust.	expertise in trafficking (n=6), homelessness/going missing service (n=4), drug and alcohol service (n=3), adult sex worker service (n=3), other voluntary sector service (n=3). **Sample characteristics:** • Age - Young people: Aged between 13 and 19. Practitioners: Not reported. • Sex - Young people: 11 women and 1 man. Practitioners: Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Young people were recruited from services that worked with young people experiencing or at risk of child sexual exploitation. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** Young people n=12 Practitioners n=90	- Lack of awareness amongst the police, or tending to associate sexual exploitation only with child prostitution, e.g. 1 police officer said: 'We don't get much here. We don't have child prostitution here. We have had the odd case. We don't have brothels here stacked with child prostitutes' (police officer, p54). - Difficulty achieving prosecutions, partly due to over-reliance on young person to press charges and give evidence. The main alternative to this was seen to be surveillance of the abuser and the young person, but noted that this is resource-intensive. 2. Social Services and ACPCs (former terminology for LSCBs) 2.1 Service response to young people at risk of sexual exploitation - systemic barriers Respondents in the study noted the following barriers to response: - Pressure on resources. - Struggling to respond to cases of known abuse, let alone tackling cases of CSE where young people did not want to co-operate. - High staff turnover. - Can be more difficult to assist older teenagers. - Lack of awareness of constrained consent/decision-making occurring in cases of CSE. - Difficulties in engaging with young people - this was sometimes seem as a failure on the part of the young person, although respondents also noted that availability and accessibility of services was also a factor. 2.2 Promising practice The study notes that promising practices included:	awareness of, and practice in relation to, CSE has changed considerably since that time. Relatively sparse reporting of interviews with children and young people.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		- Systematically identifying CSE across cases, with 1 person having the overview. - Multi-agency partnerships, e.g. between children's social care and housing. - Supporting specialist sexual exploitation services. 3. Service responses - health, education and other statutory services 3.1 Health - service provision for young people at risk of sexual exploitation Practitioners identified the need for therapeutic support for young people who have experienced CSE, although acknowledged that it can be difficult to get young people to engage. Interviewees also noted difficulties in accessing CAMHS. 3.2 Education - service provision Study notes importance of a flexible response, and enabling children to stay within school where possible. Use of learning mentors was highlighted as a something that works well. 3.3 Youth Offending Teams Those interviewed had not identified young people at risk of sexual exploitation. 4. Service response - voluntary sector specialist services Practitioners interviewed identified the following aspects of their services that facilitated engagement: - Offering a combination of 1:1 key working and counselling, along with drop-in support and group work. - Use of 'assertive outreach' for those reluctant to engage.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Offering a high level of confidentiality - Making the centre a safe and welcoming space. - Making other services accessible on-site, e.g. sexual health provision and counselling. - Offering a range of services, to give young people a number of reasons to 'hook in' (p92). Barriers cited included the service having to stop at age 18. 5. Supporting young people who have arrived from abroad - accessing protection The study notes that age problem disputes can be a barrier to trafficked and exploited young people accessing protection. Barriers to the police helping trafficked young people included lack of resources and difficulties in gathering intelligence on perpetrators. 6. Young people's views on service provision 6.1 Views on existing services The study reports that, when asked to comment on a list of workers that they found to be most helpful and supportive, most young people (n=10) chose to place their specialist young person's worker and their social worker together at the top of the list. One young person said what they valued about their social worker was that they were honest with them. Young people who thought their specialist young person's worker was the most helpful cited reasons including being able to talk comfortably and openly, flexibility and availability in emergencies. Some young	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		women valued being able to access a woman-only service. Seven young people mentioned the police, 5 'in negative terms' (p108). Schools and health services were mentioned by fewer young people, although 1 young person rated nurses to be one of the most helpful types of worker, and 2 young people had had positive experiences with sexual health nurses. Young people talked about barriers to accessing services, which included: - being reimbursed for travel expenses - language barriers - childcare. 6.2 What young people thought service provision should look like The study asked young people what they thought the ideal service would look like. Key features of what they described included: - being able to access everything in the same building - include social worker, nurses, a sexual health clinic, career advice and education support - somewhere you might go for a positive reason 'You don't just have to go there for problems' (p110) - confidential, flexible and accessible - available up to age 21 - (for young women) should be women-only - Easy to get to, with good transport links.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		One young person suggested that an online service might be helpful, particularly as it provided more anonymity. However, 2 young people rated social workers the least helpful. These young people reported that they weren't able to talk comfortably with their worker, and were not able to contact them in an emergency. One young person said: 'When I go and I say "(I am lonely)" and "I don't have a TV or radio", they say "sorry, I can't do anything". So then I keep my distance. We are put off. Social services have a limit to where they can go. You can't ask everything' (p108).	

10. Izzidien S (2008) I can't tell people what is happening at home: domestic abuse within South Asian communities - the specific needs of women, children and young people. London: NSPCC

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To consider views of managers and practitioners who work with women and children from South Asian communities in England and Wales who have been affected by domestic abuse. **Methodology:** Qualitative interviews. Semi-structured interviews with managers and practitioners (n=30) and 2 focus groups.	**Participants** **Children and young people -** Asian young girls and women (n=16). Professionals/practitioners - Focus groups and semi-structured interviews with managers and practitioners (n=30). Note, majority of professionals work for NSPCC domestic abuse services but it is not clear how many. **Sample characteristics** • Age - Children and young people - Two teen programmes with girls age 10–15 and teens for 16–19. Professionals= Not reported.	**Narrative findings** Barriers to effective response are reported in Chapter 3 titled 'Barriers to help-seeking': Chapter 3: Barriers to help-seeking: 3. 1 Recognising domestic abuse -Practitioners reported that children who were born into domestic abuse felt that it was a normal pattern of life so didn't always report incidents. One service practitioner stated: 'As a Muslim she thought that this was acceptable [for her dad to hit her mum] and was surprised to hear that it was against the teachings of the Qur'an' (p17). 3. 2 Language	**Overall assessment of external validity** + Clear relationship with guideline topic and question, however no reports on gaining ethical approval or consent from participants. **Overall assessment of credibility (internal validity)** - Not enough information on methodology.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Newham Asian Women's project conducted consultations with Asian girls and young women (n=16) through their Teens programme through two youth groups. **Country:** UK, England and Wales. **Source of funding:** Voluntary/charity. NSPCC.	• Sex - Children and young people - 16 females. Professionals - Not reported. • Ethnicity - Children and young people - South Asian communities. Professionals - Not reported. • Religion/belief -Children and young people - Not reported. Professionals= Not reported. • Disability - Children and young people - Not reported. Professionals - Not reported. • Long term health condition - Children and young people - Not reported. Professionals - Not reported. • Sexual orientation - Children and young people - Not reported. Professionals - Not reported. • Socioeconomic position - Children and young people - Not reported. Professionals - Not reported. • Type of abuse - Children and young people - Domestic abuse. Professionals - Not reported. • Looked after or adopted status -Children and young people - Not reported. Professionals - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Children and young people - Not reported. Professionals - Not reported. **Sample size -**	- Practitioners stated language wasn't always a barrier for the majority of South Asian young people because they had bilingual skills. However, it was noted by some participants that communication between schools and mothers was tenuous because in some instances 'they were only allowed out of the house to take their children to school' (p18), therefore contact was limited impacting on the ability to build relationships which impacts on seeking help. 3. 3. Lack of support from extended family - Generally practitioners and managers reported a lack of support for women and their children from extended family, it was seen as the exception, rather than the norm. One service practitioner commented that 'you will be seen as a bad woman for leaving your husband. People won't come to your house, kids can't go to parties, you will not have that social network of people coming around, you lose that connection with your own community, you are more exposed to racism. Without support it is impossible for them to leave' (p19). 3. 4 Experiences of discrimination - One service manager commented on the current political climate [post 9/11 and 7/7]: 'there is a need to engage more with South Asian women because they might be more reluctant to access services' (p19). - 1/3 of practitioners and managers cited children's issues with bullying and racism from the community. This impacts on young people feeling marginalised which might mean they mistrust professionals. - One service practitioner commented upon this marginalisation: 'One girl went home and asked her mum	**Overall score** - Relevant findings to research question, however with little methodology it is difficult to contextualise the experiences of service users. Additionally, young peoples' voice is lost in the report as only summarised at the end of Chapter 3.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Children and young people - n=16 Professionals= n=30	to bleach her face so she could be like the other girls at school' (p20). 3. 5 Insecure immigration status - Majority of managers and practitioners highlighted concern for the welfare of children whose mothers have insecure immigration status. This means that upon entering the UK, they must remain with their husband who has a secure British citizenship for 2 years, causing them to be financial dependent on their husband as they have 'no recourse to public funds'. 3. 6 Cultural barriers - The impact of shame and honour on South Asian young people was reported by practitioners and managers who spoke about the families' position within the wider context of the community. One service manager reported, 'I think the South Asian community is programmed from birth to know that there are things you don't say outside the house, and domestic abuse is one of them' (p22). - Managers and practitioners noted that there was a sense of isolation and when young people did reach out, in some instances at school, i.e. witnessing their parents arguing and fighting, they did not want the professionals to act. 3. 7 No one to turn to - 'I think the main thing is they don't speak about it. They don't have that release or outlet and nobody knows what is going on for them' (service practitioner, p25). - A couple of practitioners noted a number of case where teachers didn't believe what their Asian pupils	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		were telling them, coupled with young people not wanting the attachment of embarrassment or shame amongst their peers. 3. 8 Lack of knowledge about help available - Some managers and practitioners stated that South Asian children and young people didn't know where to turn to for support as there was a lack of information about services that offer support. - Additionally, young people were fearful of consequences, e.g. in 1 case, 'a South Asian girl wished to leave a violent household but was convinced her family would find her as the community was well connected. She felt her only option was to remain at home and put up with the violence' (p26). 3. 9 Perceived barriers: South Asian girls and young women - Of the 16 girls that were consulted, the summarised barriers to seeking help were as follows (p26): i. Many young women are trapped in violent situations due to the pressures from community. ii. People may be scared to talk because the situation could get worse. iii. Rumours spreading in the community and fears of what others may say. iv. People may not believe them or think they are stupid. v. Fear of community opinions. vi. Fear of perpetrator - threats to family and friends. vii. Trust issues and not being about to speak about abuse. viii. No services available. IX. Not aware of services. X. Fear of not being understood.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		XI. Issues around izzat and sharam for Asian families. XII. Some women may not seek help for fear of losing their children.	

11. Kazimirski A, Keogh P, Kumari V et al. (2009) Forced Marriage Prevalence and Service Response. London: Natcen

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The research had two aims: 1. To improve understanding of the prevalence of FGM 2. To examine how services are currently responding to cases of FGM. The study states that is had 'a particular focus on UK resident children and young people under 18 years of age' (p1.) We have extracted data only in relation to research question 2, which has content, which relates to our review questions 6 (Recognition), 14 (Early help) and 20 (Response). **Methodology:** Qualitative study. In-depth interviews with 40 key stakeholders across	**Participants** Professionals/practitioners - 40 professionals across 4 local authorities, covering both statutory and voluntary agencies. Respondents included: Statutory sector respondents • Police - detective inspectors, superintendents, sergeants • domestic violence (DV) - DV community safety unit (CSU) officers, DV outreach services, DV co-ordinators • child protection (CP) staff - directors of children's services, local safeguarding children's board (LSCB) co-ordinators, safeguarding children co-ordinators, CP advisors and co-ordinators • education - education welfare officers (EWOS), school counsellors, student services officers, personal advisors • local councillors • primary care trust (PCT) public health managers • housing services staff. voluntary sector respondents: • black/minority ethnic (BME) and DV - DV women's groups staff, refuge staff, counselling staff • victim support workers • law centre	**Narrative findings** These findings have been extracted from Chapter 6 on 'Case response and management'. Responses to cases of forced marriage were primarily considered to be part of domestic violence services, although responses to young people under 18 required a child protection response. Child protection responses were generally less clearly articulated that responses via the domestic violence services route. A typical response might be, after a child has gone missing from education for more than 21 days: - School writes to LA education and welfare team. - They carry out checks with housing, children's services and benefits agencies and may conduct a home visit. - If child is abroad, case is referred to Forced Marriage Unit. - If child not abroad, case dealt with by children's services. Quality and nature of response depended on the following factors:	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Good relevance to question, but no consideration of ethical issues. **Overall validity rating:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
four case study local authorities. **Country:** UK, England. **Source of funding:** Government - Department for Children, Schools and Families with support of Forced Marriage Unit.	workers • youth/children's charity workers • religious leaders. (p13). **Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** 40 individuals interviewed across 4 case study local authorities.	- Capacity of partner agencies - due to lack of resources and reported high turnover of staff in statutory children's services and schools. - Taking forced marriage seriously - priority attached to forced marriage across partner agencies. - Cultural sensitivity - there was a perception that some statutory agencies thought forced marriage was beyond their remit as it is a 'cultural issue' (p43) or considered to be a private family matter. - Compartmentalisation/culture of referral - study reports a perceived tendency for agencies to want to 'refer on' cases of forced marriage, rather than respond themselves. - Attitudes/perceptions of the victim - respondents saw part of their role as encouraging young people to recognise the risks they were facing. - Differences in partners' expertise - Respondents reported variable levels of understanding and awareness across different agencies, including awareness of what voluntary sector support services were available. - Differences in professional practices and norms - particularly between the statutory and voluntary sector. This included differences in the way that cases were drawn to the attention of services, and also the fact that the voluntary sector tended to seek solutions which maintain the family structure, which was not always possible for services operating within statutory frameworks. 6.2 Pitching the level of response The study reports that professionals found it difficult to balance being seen to take cases of forced marriage seriously (as demonstrated by taking out a care order or Forced Marriage Protection Order) compared	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		to not trying to be too 'heavy handed'. Other reported barriers to responding included the reluctance of victims to challenge their families and the fact that forced marriage may be just one factor in a more complex case (for example CSE, DV or child abuse may also be occurring). 6.3 Case co-ordination The study reports that arrangements for co-ordination of the various agencies who may be involved in a forced marriage case varied across areas, with most showing a need for better co-ordination. 6.4 Attitudes towards use of forced marriage protection orders (FMPOs) Respondents in three of the localities reported little use of FMPOs, with concerns including: that they be perceived as being 'against' a particular minority group, and that they may be perceived as having lower status than a legal response. However, police respondents in one local authority area reported making extensive use of FMPOs. 6.5 Barriers to effective case response and management The authors summarise the above barriers as follows: '• Lack of sufficient resources, especially in the voluntary sector • High staff turnover in the statutory sector, making it difficult to embed understanding and practice norms • Variability in levels of professional commitment to respond to FM • Reticence to challenge practices perceived as cultural norms	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		• Tendency to refer out of service with lack of follow-up • Limited understanding of FM • Limited understanding of the range of approaches that can be taken in response to FM, which are reliant on careful and informed risk-assessment • Lack of badly-needed case co-ordination • Gaps in service provision (specifically that for 16- to 18-year-olds) • Lack of knowledge about FMPOs • Lack of experience of working with FM leading to uncertainty around appropriate level of response and practices (causing delay in high-risk cases)' (p48). 6.6 Facilitators to effective case response and management The study identifies the following facilitators of effective case response: - Good assessment and risk assessment. - A variety of responses, ranging from counselling and support to FMPOs and prosecution where required. - All partners taking forced marriage seriously including being able to challenge practices 'without being seen to challenge the cultures within which these practices are associated' (p48), knowing when to work with a case and when to refer on, having sufficient training and practice guidelines. - Sufficient resources. - Use of FMU guidelines. - Responses to FMPOs was more mixed. 6.7 Key features of a good response The authors summarise the following features of a good response:	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		'• Individual assessment and support plan to identify services required, including employment and financial needs; • Where required, reassurance for the victim that going against FM is not going against their religion or culture ; • Encouragement and help with continuing education; • Where required, referral for counselling; • Providing information on services available, including leaflets. Where the case was deemed to be high-risk • Advice on warning signs (especially in relation to impending trips abroad); • Advice on where to hide their passport; • Taking a photograph of the young person; • Obtaining copies of passports or passport details; • Provision of a mobile phone; • The use of code words during telephone or other conversations to indicate immediate danger; • Establishing a contact able to confirm that the victim is safe should they go missing; • Establishing and agreeing measures of maintaining contact should the victim be taken out of the country; • Safety measures for trips abroad (named guarantor other than parents, contacts in the host country to check on the young person); • Flexibility in location for meetings (for example police statements and interviews taking place in a refuge if the victim is uncomfortable going to the police station); • Establishing a 'contract' with the young person stating what authority they grant the agency to intervene or enquire should they leave the country.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Where the young person had left home • Housing support • Support to encourage independence, including training on life skills • Legal advice. Other key aspects • Following up referrals (e.g. following up a faxed form with a phone call, especially in urgent cases) • Recruiting male and female workers from the local community • Seeking advice of other agencies (e.g. the FMU, specialist young person or women's group) where appropriate, which doesn't have to involve referral of case if appropriate to maintain confidentiality • Different LA departments working closely together (e.g. Social Work team attached to Homelessness Unit) • Provision of bilingual workers; • Training alongside guidelines (including on FMPOs), so professionals engage with them • Provision of drop-in advice sessions • Provision of formal 24 hour facility for reporting FM or seeking help (as some respondents reported having to give out their own mobile numbers to young people to call in the event of an emergency)' (pp49–50).	

12. McGee H, Garavan R, de Barra M et al. (2002) The SAVI report: Sexual Abuse and Violence in Ireland. Dublin: The Liffey Press in association with Dublin Rape Crisis Centre

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: Main aim of the study was to 'estimate the prevalence of various forms of sexual violence among Irish women and men across the lifespan from childhood to adulthood' (pxxxi). Additional aims included experience of services by those who had disclosed abuse - our data extraction has focused on this. **Methodology:** Survey. **Country:** Not UK. Ireland. **Source of funding** Government.	**Participants** Adult survivors of child abuse. Survey of general population, of whom 24% of male respondents and 30% of female respondents reported some form of abuse in childhood. **Sample characteristics:** In the studies for which we have conducted data extraction for this question, the sample characteristics were as follows: • Age - Note - figures relate to respondents to respondents to the survey as a whole - not just those who reported abuse. Male respondents 20-24 11.3%, 25-29 7.6%, 30-39 22.2%, 40-49 25.2%, 50-59 14.3%, 60-69 11.2%, 70-79 6.8%, 80+ 1.3%. Female respondents 20-24 6.1%, 25-29 6.4%, 30-39 24.4%, 40-49 23.3%, 50-59 20.2%, 60-69 10.1%, 70-79 7.2%, 80+ 2.3%. • Sex - 48.6% male, 51.4% female. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Social classification: Male respondents Professional 17.7%, intermediate 22.6%, routine non-manual 21.6%, skilled manual 20.0%, semi-skilled	**Views and experiences** Note: this study reports findings from people who have been sexually assaulted in childhood or in adulthood. We have data extracted findings relating to childhood abuse only. Where findings have not been disaggregated in to childhood/adulthood experiences, they are not reported as may provide a misleading picture of child experiences. **Satisfaction with services** The study examined what support services were used by those who had been sexually abused or assaulted, and survivor views on these services. The findings were as follows: • Gardai and legal system: - There were low rates of disclosure to the Gardai (4.6% of people abused in childhood). - Key reasons given for not reporting to the Gardai were: the respondents thought the case would be 'too trivial' (p131), they were too young at the time, not wanting to distress families, feeling ashamed, blaming oneself, being concerned about family reactions, thinking the Gardai couldn't do anything to help, abuse happened too long ago. People who had reported their experiences to the Gardai were asked about their experiences (note sample size relatively small). The study found that, of adults who had been abused in childhood: - 89 per cent were satisfied with how seriously the Gardai treated their situation. - 17 per cent were dissatisfied with Gardai's sensitivity to their feelings. Only 6 respondents who had been abused in childhood went on to experience court cases, and their views do not appear to be reported here.	**Overall assessment of internal validity** + Lack of disaggregation of childhood versus adulthood abuse presents a challenge in interpretation of results. **Overall assessment of external validity** ++ **Overall validity score** + Lack of disaggregation of childhood versus adulthood abuse presents a challenge in interpretation of results.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	manual 12.1%, unskilled manual 1.1%, unclassified 4.9%; Female respondents Professional 10.7%, intermediate 30.3%, routine non-manual 29.0%, skilled manual 6.9%, semi-skilled manual 19.1%, unskilled manual 0.9%, unclassified 3.1%. • Type of abuse – Abuse experienced during childhood (under 17) (wording of questions made clear that these were non-consensual experiences) • Looking at pornographic material - male 6.7%, female 2.7%. • Being photographed or videoed - male 1.0%, female 1.3%. • Someone exposing their sexual organs - male 12.5%, female 20.6%. • Someone masturbating in front of you - male 6.2%, female 5.3%. • Being touched in a sexual way - male 11.2%, female 14.9%. • Someone else get you to touch them in a sexual way - male 9.7%, female 9.0%. • Someone rub their genitals against your body in a sexual way - male 6.6%, female 10.1%.	• Medical professionals: - The study reports that only 17 respondents abused in childhood reported it to a medical professional. - Of those reporting child sexual abuse, 33% were dissatisfied with the experience, with 11% reporting that they felt that the medical professionals made them feel 'responsible for their experience of sexual violence' (p141). - 20% of respondents said they were dissatisfied with how physical examination was explained to them and conducted. • Counselling/psychological professionals: - The proportion of respondents who had been abused as children and were satisfied with their experience of counselling is not reported. • Health Board Services: - Only two participants had experiences of Health Board Services (unclear if they had been abused in childhood or adulthood). **Experiences of marginalised subgroups - satisfaction with services** • Homeless women and their children: Findings largely appear to relate to sexual violence experienced in adulthood - not disaggregated. • Travellers: Additional detail relating to disclosure: The study notes that respondents from the Traveller community described a culture in which disclosing abuse was seen to bring shame and dishonour on your family. It was noted that this would be particularly severe for boys in the community. This linked to experiences of involving services. For example, one participant said: “Traveller women would be very afraid to speak to a social worker. They'd be put on the 'at risk' register,	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Someone attempt to have sexual intercourse with you - male 3.0%, female 4.6%. • Someone succeed in having sexual intercourse with you - male 1.1%, female 1.7%. • Persuaded to have oral sex - male 1.1%, female 0.9% 11. • Persuaded to have anal sex - male 0.9%, female 0.3%. • Insertion of fingers or objects in to vagina or anus - male 0.6%, female 4.4%. • Looked after or adopted status – Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** Sample size - 3120 respondents.	wouldn't they" (p209). Study notes that, for this group, further practical barriers to accessing services can include 'isolation, illiteracy and a lack of information, money and transport' (p210). • Prison population: Relates to experiences in adulthood only. • Women in prostitution: Survey completed by staff and volunteers only. • People with learning disabilities: No primary data gathered. • Psychiatric inpatients: Research with staff only.	

13. McNaughton Nicholls C, Harvey S, Paskell C (2014) Gendered perceptions: what professionals say about the sexual exploitation of boys and young men in the UK. London: Barnardo's

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The study has 4 research questions, 1 of which matches our review question which is: - to 'suggest ways in which policy and practice	**Participants:** Professionals/practitioners - Professionals with experience of working with boys and young men experiencing, or at risk of, sexual exploitation.	**Narrative findings** 3.4 Effective responses The study reports that professionals held differing views on whether or not specialist gender-based ser-	**Overall assessment of internal validity:** + **Overall assessment of external validity:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
may be able to identify and appropriately respond to male victims of CSE, as well as those at risk' (p13). The other 3 questions are less relevant to this review question which are: - identify perpetration and victimisation processes apparent in male-victim CSE cases known to professionals - explore existing service provision for boys and young men at risk of or experiencing CSE - identify future research priorities (p13). **Methodology:** Qualitative study. This paper reports a qualitative study. It appears that this was undertaken as part of a wider study (summary reported in McNaughton Nicholls et al. 2014 'Research on the sexual exploitation of boys and young men'). **Country**: UK, England.	**Sample characteristics:** • Age - Not reported. • Sex - Female: n=29 Male: n=21. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** n=50, comprising 41 qualitative interviews and 9 'online responses' - unclear what the online responses involved.	vices were required. However, those already providing specialist services for young men argued that these services not only supported young men, but also raised awareness amongst other professionals. 3.4.1 Gender differences in support needs Professionals thought that boys and young men had gender-specific support needs, related to their barriers to disclosure. The perceived specific needs included: - Sexual identity: Support for gay and bisexual young men regarding sexual orientation; support for young men identifying as heterosexual who have experienced same sex abuse; support regarding healthy sex and relationships. - Expressions of masculinity: Support for understanding different representations of masculinity; healthy male role models. - Psychological needs: Support to acknowledge exploitation; understanding of different methods of self-harm; support regarding anger - Criminal involvement: Support to understand criminality as a response to trauma. 3.4.2 Engaging boys and young men with services Professionals identified principles for effective engagement regardless of gender. These included: - a positive consistent relationship - feeling listened to and respected - being empowered to make their own decisions - having flexibility in how they engaged with services - being able to stay engaged with services over an extended time period if required.	UK study but only part of overall research aim was relevant to our review question. **Overall validity rating:** + Only part of overall research aim was relevant to our review question. Study is of reasonable quality, although limited exploration of divergent perspectives across different types of interviewees.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Source of funding: Voluntary/charity - The Nuffield Foundation.		3.4.3 Gender of support worker Most professionals thought that both men and women could effectively work with male service users. However, it was thought important that a young person had a choice about the gender of their worker, particularly when working with trans* young people. 3.4.4 Ways of working with young men Two main approaches are reported. Professionals who attributed gender difference to socialisation thought that young men needed to be supported to challenge stereotypes and talk about their experiences. Professionals who attributed gender differences to innate biological differences placed more emphasis on allowing young men to express themselves in different ways, for example through activities. Practitioners suggested that there was a risk of reinforcing gender stereotypes through modes of engagement, for example there might be a greater tendency to focus on activities about 'moving on' with young men compared to young women. One practitioner said: 'Society actually criticises men for not being able to show their emotions and, you know, "boys don't cry" and then […] the services just perpetuate that […] because it's like "Let's just go in there, you know, let's just move on, get a job, everything will be fine. Go to your training and get to college"' (practitioner, CSE service, p40).	

14. Pearce J (2011) Working with Trafficked Children and Young People: Complexities in Practice. British Journal of Social Work 41: 1424–40

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The research aimed to: - explore the different ways that 'trafficking' is understood by a range of practitioners from different service agencies; - look at the obstacles that emerge when trying to identify trafficked young people; - chart the process through which a child or young person first gained access to a support agency; and - identify how the practitioner understood the immediate and longer-term needs of the children and young people concerned (p1427). **Methodology:** Qualitative study. A total of 72 practitioners were interviewed by individual interview or in a focus group (n=9 focus groups). **Country:** UK. Three locations in England.	**Participants:** Professionals/practitioners - Total 72 practitioners: Social workers - n=22 specialist children's NGOs and separated children/asylum workers - n=12 Police/CPS/YoT/Border Agency - n=11 Residential child care and statutory children's centre workers - n=10 Health workers - n=10 Education workers - n=7. **Sample characteristics:** • Age - Not reported. • Sex – Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** Total of 72 practitioners took part - 3 focus groups were run in each of the three areas (n=65) - Semi-structured	**Narrative findings** Factors that hinder effective response: - Practitioners understanding of trafficking was seen as varied. One participant commented, 'I have looked at all the different definitions that they have and I realised that there isn't a full definition that everyone sticks to. It can be looked at very differently' (p1428). - Some suggested that they had theoretical knowledge of trafficking but struggled to apply in practice due to limited experience. - The contention between whether a child can consent to being trafficked, i.e. 'willingly trafficked', confused practitioners which can result in the child being overlooked. - The authors coined the term 'culture of disbelief' which is where practitioners who are unaware of indicators of trafficking and find it difficult to believe a child has been trafficked. One professional commented 'a cynic might say she was trying to get money out of us ... she was pregnant at the time' (p1431). Factors that help effective response: - Professionals commented on taking a child-centred approach to their work in order to alleviate a sense of responsibility on the child. One participant commented how they used similar language to the child and listened to them: '... they transfer the care if a person who suspected trafficked them to the care of Social Services ... they are then able to compare their life before hand and their life now, and they can identify that they have been maltreated. Have I ever heard	**Overall assessment of internal validity:** ++ Excellent, thorough study that meets aims and objectives. Authors have cited other studies to discuss the current complexities working with trafficked children through interviewing 72 practitioners. **Overall assessment of external validity:** ++ Relevant study to research question with excellent ethical consideration. **Overall validity rating:** ++ Most criteria is met and very comprehensive, in-depth findings.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Source of funding: Voluntary/charity - NSPCC funded the project.	interviews with 7 practitioners from each area.	a young person use the word “trafficked”? No never’ (p1432).	

15. Pearce J, Hynes P, Bovarnick S (2009) Breaking the wall of silence: practitioners’ responses to trafficked children and young people. London: NSPCC

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The aims of the research are as follows: 1. Explore in depth the different ways in which trafficking is understood by a range of practitioners from different service agencies and provide evidenced recommendations for practice in their area. 2. Explore the obstacles that might emerge to identifying the numbers of young people trafficked in the 3 areas. 3. Identify the numbers of children and young people trafficked into each of the 3 areas. 4. Chart the process through which a child or young person first	**Participants** **Sample characteristics** • Age - Practitioners in focus groups/interviews: Age not reported. Young people’s case files: Age 3 and under n=6, age 4-8 n=1, age 9–12 n=1, age 13–15 n=15, age 16/17 n=14. • Sex - Practitioners in focus groups/interviews: Gender not reported. Young people’s case files: Girls n=30, Boys n=4, Gender not known n=3. • Ethnicity - Practitioners in focus groups/interviews: Ethnicity not reported. Young people’s case files: Ethnicity not reported, but information on nationality provided. Country of origin: UK n=1-, China n=8, Nigeria n=8, Somalia n=1, Pakistan n=1, Cameroon n=1, Ghana n=1, Congo n=1, Sierra Leone n=1, Zimbabwe n=1, Uganda n=1, Eastern European Country n=1, Unknown n=1.	**Narrative findings** 1. Responding to trafficking: the role of mainstream services The study reports that practitioners thought that mainstream services had an important role to play in protecting and promoting the wellbeing of trafficked children. Practitioners also noted that: - Multi-agency work is crucial, supported by the LSCB and local arrangements such as local protocols - It is important to have co-ordinated information sharing and joint work between police and child protection workers - Practitioners needed a good awareness of indicators of trafficking, which may include criminal behaviour - Young people needed to be supported to access mainstream health provision, and mental health provision, and to stay in education - Young people may need additional support at ages 16 to 18 when their legal status in the UK may start to come into question.	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** ++ **Overall validity rating:** ++ Thorough data collection, analysis and reporting.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
gained access to a support agency, including how they first contacted an agency and for what reason. 5. Where possible, provide a profile on each of the children and young people identified including: age; nationality; country of origin; the reason they were trafficked into the country; and a summary of their current circumstances. 6. Identify how the practitioner understood the immediate and longer-term needs of the children and young people concerned. 7. Identify how the professionals feel these needs are best met. 8. Where possible, identify perceptions of how the children/young people feel these needs are best met. 9. Make recommendations about how agencies or individuals can best support the children/young people	• Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Of the case files examined, 10 related to cases of trafficking of UK citizens and 27 to cases of trafficking in to the UK from abroad. Reasons for trafficking were as follows: Sexual exploitation n=19 (this included 9 of the trafficked UK citizen), benefit fraud/illegal adoption n=7, domestic servitude n=5, forced marriage n=2, restaurant work n=2, drug trafficking n=1, not known n=1. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - All young people were trafficked. **Sample size** Practitioners: n=72 Children's case files: n=37.	2. Responding to trafficking: specialist services The study reports that practitioners advocated three types of specialist service: - Trained and specialist interpreters - Safe and supported accommodation with trained foster carers - A dedicated keyworker approach to service delivery. The study notes that practitioners highlighted the importance of interpreters who had been trained to understand and manage that young people's accounts of trafficking may be affected by ongoing threats from their traffickers. Practitioners also thought that existing local authority accommodation was not well equipped to support trafficked children, including the availability of emergency placements for those who have just arrived in the country.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
concerned. We considered questions 2 to be relevant to our review question on Recognition, and 7 and 8 to be relevant to our review question on Response. **Methodology:** Qualitative study. - Focus groups with 65 practitioners. - Interviews with a selection of focus group practitioners (number not specified) and with an additional 7 practitioners - Case file analysis of cases of 37 trafficked children and young people. **Country:** UK, England. **Source of funding:** Voluntary/charity - Study reports that research has been funded by 'The Children's Charity' (p6) (unclear if this refers to NSPCC or another charity).			

16. **Rees G, Gorin S, Jobe A et al. (2010) Safeguarding young people: Responding to young people 11 to 17 who are maltreated. London: The Children's Society**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The aim of this study is to explore 'access to, and initial responses of, services for young people with potential maltreatment ... to promote protective responses for this target group' (p7). The section relevant to this review question is entitled young peoples' experiences of seeking help. **Methodology:** Qualitative study. The relevant methodology relating to this scope involved in-depth interviews with 24 young people who had been referred to children's social care aged 11–17. This study is also reported in Jobe and Gorin (2013). This paper reports a briefer version of the study findings, but has more detail on study methods. Where necessary,	**Participants** Children and young people - The 24 young people who were interviewed for our study either had social care intervention from an early age or had first come to the attention of Children's Social Care Services in between the ages of 11 and 18. **Sample characteristics** • Age - Young people ranged between 11 and 18, with categories determined: 11–14 (n=5); 15–16 (n=13); and 17–18 (n=6). • Sex - The study includes 14 males and 10 females. • Ethnicity - Participants were White British majority (n=18). One young person was British Asian and the study included unaccompanied asylum seeking children who were originally from Afghanistan (n=3) and Eritrea (n=2). • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - The maltreatment experience was divided into two	**Narrative findings** Findings are titled 'Young people's experiences of contact with children's social care' and details factors that help and hinder effective response: 1. Young people's relationship with social work professionals: - For young people, the relationship with the social worker was considered pivotal, 1 based on feeling they were being listened to and there was time to build effective relations. One young person, aged 15 cited that her social worker is '... really really nice ... Like she's really easy to talk to and really chatty. She's a lot more helpful than the first one, like I've had regular meetings with her, and we've done like mind maps of family and like putting people who are closer in the inner circles and stuff like that' (p52). - UASC's commented positively upon building a consistent relationship with their social worker in a specialist unaccompanied minors team – e.g. 1 young person aged 17 stated: 'They help me do everything ... everything, everything. When I come to [social services office] when I have an appointment ... I just go reception "I want to speak to my social worker" - he will call her [and she is] coming down ... I like social services, they're really nice' (p53). - Conversely, some young people had established negative relationships, mainly due to the inconsistency of their worker and being assigned a new one. In addition, retelling their story was difficult to speak about and relive difficult memories.	**Overall assessment of internal validity:** + (When taking in to account additional info from Jobe and Gorin) The study does not have a rigorous methodology or consideration of limitations. Presentation of information is difficult to ascertain where data is collected making conclusions challenging to draw. In addition, there is discrepancy in young people's age as referred in text to both: 11–17; and 11–18. **Overall assessment of external validity:** ++ Study relates to question of exploring young peoples' views and experiences of response from children's services and what helped/hindered effective support.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
additional methodological information has been taken from Jobe and Gorin (2013). **Country:** UK. This study is based in UK.	groups – those who had suffered maltreatment from an early age therefore were receiving social work intervention from an early age but were still receiving support between the population age of 11–17 (n=6). The majority came to the attention of children social care (n=18). Reason for referral included a range of issues such as 'homelessness, being thrown of home, mental-health problems, alcohol and drug misuse, behavioural problems, risk-taking behaviour, violence and conflict with parents' (p39). • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - The study included unaccompanied asylum seeking children who were originally from Afghanistan (n=3) and Eritrea (n=2). **Sample size** Young people – n=24.	- Some young people also described their frustration at having infrequent meetings with their social worker, i.e. not getting contact details or if their worker was constantly unavailable. One young person remarked: 'Sometimes when I ring [my social worker] she never rings us back' (p55). 2. Obstacles to building positive relationships 2.1 Lack of clarity - A number of young people expressed confusion over the professionals role in the safeguarding process, 1 young person commented '... to be honest having a social worker king of confused me a bit, she was asking all these complicated questions and I was 11 at the time, thinking, what? What's that mean? (Laughs) Really confusing' (p56). - Children and young people that were Looked After appeared to have a clearer idea of child protection process, than those respondents who had a shorter social care involvement. 2.2 Being listened to an informed - Some young people felt their accounts were not taken into consideration which undermined their confidence. As one young person recalls her experience being taken into A&E after she was physically attacked by her father, the medical staff contacted children's social care: 'No, the social worker just came to the emergency room ... I was kind of confused ... I didn't talk to them, they just talked to my mum and my dad and that was it' (p57). 2.3 Managing expectations	**Overall validity rating:** + The study is suitable for scope and the findings enrich discussion about barriers to young people disclosing sexual abuse. Drawing on additional information from Jobe and Gorin (2013), where the research design is more informed, the findings are more convincing as data is richer and analysis is clearer.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		- Young people described feeling let down if they were given 'false expectations by professionals about what might happen' (p57). 2.4 Young people not being given a say/autonomy - In some instances, there was a tension between meeting the needs of the child and the parents, which caused young people to feel upset if their parent's views were taken into consideration above theirs. As noted, 1 young person who had built a positive relationship with a social worker was changed due to her mum complaining about her. - Confidentiality was raised as an issue especially if the young person felt the social worker would tell the parents what was discussed. 3. Social work responses - Some young people felt that children's social care did not react fast enough to protect them. In the 3 case studies explored in the paper, age was highlighted as a potential factor in them not being considered a priority. 4. Experiences of child protection conferences and looked after children reviews - A number of young people described these meetings as difficult to engage with and they felt their voice was lost. - In addition, not knowing all the professionals that were present was a challenge and didn't always make the young people feel comfortable. For example, one young person commented, 'It's just like the chairperson that comes, I just don't know them and I'm like "Well can you not keep the same chairperson to like handle the meetings, so like I actually know them?"	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Because it's not very good somebody coming into the meeting and you ... and you're discussing your personal issues with everyone, bar one person that you don't know. I mean I know they're not going to say anything, but it's just I feel ... well I don't really want them to know because I don't really know them' (p63). - On the other hand, other young people felt listened to in their LAC Reviews: 'I enjoy going to them cos I get my say really' (p63).	

17. Richardson Foster H, Stanley N, Miller P et al. (2012). Police intervention in domestic violence incidents where children are present: police and children's perspectives. Policing & Society 22: 220–34

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To assess views and experiences of children on police intervention in domestic violence and police professional practice in response to domestic violence. **Methodology:** Qualitative study. This study was a mixed methods study examining quantitative data (prevalence and nature of domestic abuse) as well as qualitative data on views and experiences of children on police intervention in domestic violence and police	**Participants** Children and young people – n=19 Professionals/practitioners – Police officers n=30 **Sample characteristics** • Age - Children and young people; aged 10–19 years Police officers: not reported • Sex - Children and young people: n=19 (8 males and 11 females). Police officers: n=33 (no info on gender). • Ethnicity - Children and young people: 16 White British, 1 as White and Asian, 1 as White and Black Caribbean, 1 as White and Black African. Police officers: not reported. • Religion/belief - Children and young people: Not reported; police officers: Not relevant.	**Narrative findings** A. Young people's attitudes to and experiences of the police: 1. Negative experiences of the police - cynical and distrustful attitudes about the police generally, described as 'biased', 'judgemental' and 'ignorant'. 2. Young people felt the police focus was on the adults and were themselves frequently excluded by the police. They emphasised the importance of officers listening to their accounts and validating the seriousness of their experiences, like being acknowledged, heard and believed both at the time of the incident and in their subsequent interactions with the police. 3. Young people considered that their accounts lacked credibility in the eyes of the police, they wanted explanations from the police about what would happen next, in particular, they wanted to know whether it was likely the perpetrator would or could	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
professional practice in response to domestic violence. Only the qualitative data on views and experiences were data extracted to answer this review question. **Country:** UK: 2 sites in northern and southern England. **Source of funding:** Not reported.	• Disability - Children and young people: Not reported; police officers: Not relevant. • Long term health condition - Children and young people: Not reported; police officers: Not relevant. • Sexual orientation - Children and young people: Not reported; police officers: Not relevant. • Socioeconomic position - Children and young people: Not reported; police officers: Not relevant. • Type of abuse - Children and young people who witnessed domestic violence. • Looked after or adopted status - Children and young people: Not reported; police officers: Not relevant. • Unaccompanied asylum seeking, refugee or trafficked children - Children and young people: Not reported; police officers: Not relevant. **Sample size** Children and young people: n=19 Police officers: n=30	return. They also felt that their opinions were not always considered when decisions were made about the next steps for the family. 4. Young people felt positive when being spoken to directly and offered support by the police. Response times by the police 5. Young people emphasised their need for a swift response and a slow response was interpreted as reflecting the lack of priority assigned to domestic violence. This dissatisfaction with response times appeared to be underpinned by concerns about whether the police were taking children's accounts seriously and validating their positions as victims of domestic violence. B. Police perspectives on communication with children: 1. Frontline police officers confirmed a limited engagement with children and young people and expressed divergent views on their role, e.g., whether it was within their remit to check the house for the presence of children and to check on their welfare at the scene of a domestic violence incident. 2. Some police officers expressed reservations and reluctance about speaking to children directly, that it was not their role to speak to children at incidents, due mainly to a lack of confidence or skills in talking to children. They considered that expertise in talking to children resided with those officers who worked in specialist child protection posts. 3. They felt the reluctance to explain and speak to children directly at the first response was mainly the concern of evoking a conflict of loyalties or distress in	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		children (between parents and children and between siblings). 4. Frontline officers were anxious about opening up areas of emotional need in children as they felt that they were not equipped with relevant information about support groups or sources of help to offer children and young people. This reflected a gap in service provision. 5. A reluctance to engage with children's needs in the absence of resources was reinforced by operational procedures that militated against in-depth exploration of need, e.g., performance targets ensuring a rapid response as a priority to a domestic violence call could lead to them dashing from one incident to the next with little opportunity to engage with the individuals involved, missing the potential for communication with the child. 6. Frontline police officers considered children as neither the victim not the perpetrator so were not a primary focus for police attention at an incident and the prospect of talking to them as 'dragging them into it' or 'bringing them in'. 7. Generally, police viewed children and young people as figures on the sidelines of domestic violence incidents. In denying children a role in the experience of domestic violence, police officers ran the risk of colluding with parental claims that children were unaware of and unaffected by such violence.	

18. Skinner T, Taylor H (2009) 'Being Shut Out in the Dark'. Feminist Criminology 4(2): 130–50

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: Due to the changes in police responses to survivors of	**Participants:** Children and young people - 9 young people were interviewed	**Narrative findings**	**Overall assessment of internal validity:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
rape and sexual assault, this study seeks to understand the experiences of young people who reported a sexual offence to 1 police service in England (n=9). **Methodology:** Qualitative study. This study is part of a larger evaluation of a Sexual Assault Referral Centre (SARC) young persons' project. Qualitative face-to-face semi-structured interviews were conducted with 9 young people aged 14–16 through the SARC. Participants gave consent and 5 young people were accompanied by their mother. The mothers who attended were interviewed too and 1 father was contacted by telephone. Additionally, the research team collected feedback on the research process. Young people were	Caregivers and families - 5 mothers who attended their child's interview were also interviewed, as was 1 father via telephone. **Sample characteristics:** • Age - Young people: 14 years old (n=1); 15 years old (n=4); and 16 years old (n=4). Caregivers and families: Not reported. • Sex - Young people - 9 females. Caregivers and families: 5 females and 1 male. • Ethnicity - Young people: The authors are not explicit that young people are white but allude to the fact that no BME survivors were interviewed. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - All suffered a rape or sexual assault. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size:** Young people: n=9	Findings are illustrated by the disclosure journey of a young person reporting sexual assault or rape. The author has included 1. Initial Contact with the police and the initial statement; 2. Forensic medical examination; 3. Giving a formal video statement to the police; 4. Follow up interviews; and 5. Information about the status of the case: 1. Initial Contact with the police and the Initial Statement: - Most young people (n=7) reacted positively to the disclosure to the police with factors helping being: female officers; plain clothes and unmarked police car; and the interviewing technique e.g., 1 young person said 'They asked me if I - if I wanted to carry on' (p137). However 2 young people commented on having an uncomfortable experience because the police officer asked about her sexual activity which made her feel 'spoke down' to. 2. Forensic medical examination: - Most participants (n=8) had a forensic medical examination and the findings were: lack of choice about having a medical; lack of choice about sex of the doctor; a sense of control during the medical examination; concern about spending time waiting before the examination; and how comfortable the examination and video suites were. - Generally, participants commented that the loss of choice about whether to have examination or not was negative, however there was a sense of necessity. - 5/8 participants were seen by a male doctor, and this was concerning to most participants who commented that they felt uncomfortable not having a	The study postulates that there are limitation in the small-scale study (n=9) and its generalisability, however the experiences shine a spotlight on hard to reach groups such as young people who have been sexually assaulted. The sample is from a SARC who has supported the survivors which opens the study to potential bias. **Overall assessment of external validity:** + Overall, study meets most of the quality criteria however caution to generalise the UK as the study is based in one area where participants are collected from a Sexual Assault Referral Centre and this is written 13 years ago so unsure if SARC still exist. **Overall validity rating:**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
asked and given details about additional support from the SARC or another agency. **Country:** UK. Interviews were sought through a SARC in England. **Source of funding** Government - Home Office.	Caregivers and families: n=6	choice over the sex of their doctor despite them being described as 'gentle' and 'really nice' (p139). 3. Giving a formal video statement to the police: - Not relevant 4. Follow-up interviews: - One participant commented that the experience of follow up interviews with the police was lengthy. - Another participant reported that 'They [the police] asked me then "Did it really happen?" So, I couldn't believe it when they said that, I was proper devastated when they said that and then, then after they found out that [I was telling the truth] they did actually apologise'. - However, one participant said that she felt believed and appreciated this. 5. Information about the status of the case: - Responses varied about the criminal justice process and the information that was received. In some instances participants understood the process because it was explained to them. - However a couple of participants said that they felt 'shut out in the dark' (p142), and once the police had the information they had little contact. - SARC provide a case-tracking service for survivors keeping them informed through the criminal justice system, and some parents were positive about this. One mother commented 'Any time the lads [alleged offenders] went in, they would call back and they wrote a letter saying that they had been released on bail and their case was still pending and the they would have go back to the police station on such and such a date. And they were really spot on!' (p132).	+ Findings are relevant in part to the research question, however as small scale study there is a limitation in generalising conclusions. The authors' discourse suggests that young people have little participation or choice throughout the disclosure journey and are not informed about the process enough so makes recommendations on the findings presented in the study to be more supportive and inclusive. Furthermore, there is no information about the analysis of findings.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		- Other parents were angered having to chase information and in some cases, a survivor found out the alleged offender had been released by a friend of a friend or the local newspaper. - Eight of the participants or their parents had to contact the police or SARC to get the information about the case. - Some survivors commented on the police officer in charge of their case and they stated their experience was that they were never available or would get back to them. - The lack of information about their case was considered to be negative especially coupled with a negative outcome in court. - Consequently some participants commented that they would not trust the police again and that despite reporting being 'the right thing to do ... it seemed like a waste of time' (p144).	

19. Smeaton E (2013) Running from hate to what you think is love: the relationship between running away and sexual exploitation. Ilford: Barnardo's and Paradigm Research

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The overall study aims are to: 'collect data relating to the experiences of young people under the age of 16 who experience both running away and CSE - collect data from practitioners and projects working with young people who experience	**Participants** Children and young people - 41 young people with experience of running away and CSE whist under 16. Adult survivors of child abuse - 12 of the young people involved were over the age of 18. Professionals/practitioners - 28 projects working with young people experiencing CSE and/or running away via survey; 27 professionals working with young people with experience of	**Narrative findings** The study reports that the research with professionals and young people identified a range of factors that facilitate and hinder meeting the needs of young people experiencing running away and CSE. 4.1 Resource issues Professionals recommended that there would be more funding available for work with young people running away and experiencing CSE. Professionals	**Overall assessment of external validity** - **Overall assessment of internal validity** + Could be more detailed description of how addressed ethical issues. **Overall score**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
both running away and CSE - produce an evidence-base that outlines the relationship between running away and CSE and supports recommendations to support policy and practice responses to young people who experience both running away and CSE - produce a final report outlining findings, a summary document and a tool-kit for practitioners - work with key national agencies to ensure evidence-based findings are incorporated into national policy and practice' (p11). The interviews with young people focused on: 'a history of the young person's life and events and experiences they considered to be important - experiences of running away and CSE - what could have prevented them from experiencing both running away	running away and CSE via telephone interview. **Sample characteristics** • Age - Young people: 14 n=4 15 n=8 16 n=8 17 n=9 Adult survivors 18 n=7 19 n=2 20 n=2 21 n=1 Professionals - not reported. • Sex - Young people: Female n=25 Male n=15 Transgender n=1 Professionals: Not reported. • Ethnicity - Young people: White British n=32 Mixed Black Caribbean/White British n=3 Mixed Asian/White British n=2 Roma Traveller n=2 Bengali n=1 Sikh n=1 Professionals: Not reported. • Religion/belief - Not reported. • Disability - Young people: Self-defined learning disability or difficulty n=17, this comprised SEN n=9, general learning difficulties n=4, Attention Deficit Hyperactivity Disorder (ADHD) n=2, dyslexia and ADHD n=1, dyspraxia n=1. Professionals: Not reported. • Long term health condition - Not reported. • Sexual orientation - Heterosexual n=29, 'self-defined as gay' (p13) n=10, bisexual n=1, uncertain about their sexuality n=1. Professionals: Not reported. • Socioeconomic position - Not reported.	identified that practice was facilitated by use of voluntary funds, rather than when money was strictly ring-fenced for particular purposes. 4.1.1 Resource issues hindering meeting the needs of young people who experience both running away and CSE Research participants identified the following issues: - Funding cuts, 'predominance of short-term funding cycles' (p59). - Specialised projects cannot meet demand. - Lack of services for young people who experience both running away and CSE, particularly in rural areas. - Lack of appropriate supported accommodation, and use of out of area placements. One professional said: 'We've loaded these children's homes with young people who are at risk of sexual exploitation and it actually destabilise that home so that we can't use it anymore. And then you get young people running away together so they actually strengthen their networks' (p60). Lack of therapeutic accommodation. 4.2 Factors relating to multi-agency approaches to running away and CSE 4.2.1 Factors facilitating general multi-agency working The study identified the following factors: - Effective working relationships with other local voluntary agencies. - Strong relationships with the police. - Effective working relationships with schools. - Working with health professionals and sexual health clinics.	- Survey of services is entirely of voluntary sector services, and it is unclear whether interviewed professionals represented a wider range of services - the voluntary sector perspective of the research is not highlighted or justified in the research methodology. Little consideration in the findings of how contextual and demographic factors shape participant responses.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
and CSE - their experiences of support seeking - recommendations to both prevent and respond to running away and CSE' (p12). The consultation with professionals focused on: the nature of the relationship between running away and CSE 'identification of factors that facilitate projects' work with young people who experience both running away and CSE - identification of factors that hinder projects' work with young people who experience both running away and CSE - identification of groups of young people who experience both running away and CSE that projects find difficult to engage - gaps in national and local policy to meet the needs of young people who experience both running away and CSE - gaps in national and local practice to meet the needs of young	• Type of abuse - Young people had experienced sexual exploitation. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** Young people: n=41 Professionals via survey: 28 projects Professionals via telephone interview: n=27.	- Having health workers based within specialised projects, e.g. having a CAMHS nurse based in a specialist CSE project. - Engagement with A&E departments. - Good relationships with individual social workers. - Having co-located teams with designated workers. - Contributing to Multi-Agency Risk Assessment Conferences (MARACs). 4.2.2 Factors hindering multi-agency working - Lack of support from Local Safeguarding Children Boards - Clash of working cultures between the voluntary and statutory sectors. 4.3 Factors relating to collating and sharing data and information The study reports that professionals noted the importance of collecting and sharing information, particularly when young people move across areas, and that the failure to do this in some places hindered responses to young people who experience running away and CSE. Professionals noted that: - attitudes to sharing information were important in supporting the work of specialist services - agencies may have varying approaches to information sharing - specialist projects can be a good source of information - missing person reports are also a good way to ensure that children who are experiencing exploitation AND running away are identified.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
people who experience both running away and CSE' (p14). Here we have data extraction information which is relevant to: young people's views on recognition - aspects of professional practice which help/hinder recognition - young people's views on professional responses - aspects of professional practice which help/hinder response to CSE. **Methodology:** Mixed methods. Study comprises: - Interviews with 41 young people who had experienced both running away and CSE whilst under the age of 16 - A survey of 28 projects working with young people experiencing CSE - Telephone interviews with 27 professionals who work with children who experience running		One professional said: 'So when I'm made aware of a young person who may need the CSE service, the first thing I will do is look at the MISPER reports and see how many times that young person has been reported missing … Because we're getting those MISPER reports on a weekly basis, we can really map out and track someone's progress, someone's deterioration and there's been times when … we've been able to say "well … that boy can really do with the X [the CSE] service" and no one's flagging that to us and then I go out to the people involved with that young person, often social care, and say "do you want to refer them to us?". I will say, for example, that not only have they been missing ten times but they have been found at inappropriate addresses, they're found in the red light area, etc. etc. So we can proactively target young people at risk' (p68). 4.4 Professional awareness and knowledge The professional interviewed thought that: - There was a general lack of awareness amongst the statutory sector of running away and CSE, including the perception that running away and CSE is a 'lifestyle choice' (p69). - The concept of 'constrained choice' was useful, which states that 'young people's lack of power relating to age, need and social vulnerability also makes it impossible to give their consent to being sexually exploited' (p69). - Other professionals may be less keen to respond to older children, such as those aged 16 and over. - There is a need for greater awareness about policy and the law in relation to CSE, and of raising general awareness about the issue.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
away and CSE. The methods section also mentions the development of, and work with, a Research Dissemination Group. However the activities of this group are not reported here. **Country:** UK, England. **Source of funding** Voluntary/charity - Comic Relief.		- Use of language is important, because young people might not see themselves as having run away they have just 'stopped out' (p71). Young people also emphasised the importance of raising awareness amongst young people and for professionals, through training. 4.5 Factors relating to the local authority Issues not relevant to aspects of professional practice - relates to organisational factors within the local authority. 4.6 Factors relating to the criminal justice system 4.6.1 Young people's experience of being part of a police investigation into CSE Young people who had experienced being part of a police investigation into CSE found this stressful and difficult. This perspective was also emphasised by professionals. 4.6.2 Factors relating to the police Young people: - Emphasised that the police should provide an 'appropriate' (p76) response. One young person said: 'Don't be judgemental when you [police officers] first meet the young person like some police officers when they first met me … [Some] would make a judgement straight away after meeting me' (p76). - Suggested that the police should give more thought to why they have run away, and that the best option for them may not be to be returned to where they have come from.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		- Thought the police did not always take appropriate action against perpetrators of CSE, e.g. 1 young person said of an older male perpetrator: 'They never seized his phone or anything and he was grooming me for nine months and I never understood why they [the police] never arrested him … He [the lead police officer on the young person's case] was like "I'm too busy with other cases [against other perpetrators of CSE] you've given me" and I was like "that's no excuse; he could still hurt somebody else [another child or young person]. Why aren't you arresting him?" I got into a massive fight with the police about that and they still didn't do anything so they [the police] just need to take it [sexual exploitation] more seriously' (p77). Professionals: - Noted the importance of the police in responding to CSE, and in having good relationships between specialist CSE projects and the police. However, it was noted that the police response can be variable. Some professionals thought that the police could give 'mixed messages' (p76) about the importance of reporting young people as missing. One professional said 'Parents and carers of young people who are involved in CSE are being told by all professionals about the importance of reporting their child as missing but when they go to the police and they report their child as missing, they're being told that they [the police] are not a taxi service, have you actually looked for the girl – they're only two hours late. You know, they [parents and carers] could be at a child protection conference where someone says "why didn't you report your child [to the police] whilst she was missing? Well, I tried to do that but the police told me I	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		was wasting their time" … Parents can be seen to be not engaging or failing to protect their children when actually they're not but they're getting the wrong response [from the police]' (pp76–7). 4.6.3 Criminalisation of young people who experience running away and CSE Professionals noted that young people involved in CSE are often 'criminalised' (p78). One young person also gave an example of this: 'I had a standoff with the police with a knife. For about half an hour I held it [the knife] to myself and was saying that I was gonna hurt myself. They [the police] took it [the knife] off me and arrested me for having an offensive weapon' (p78). 4.7 Factors relating to specialised projects Issues not relevant to aspects of professional practice - relates to organisational factors within voluntary organisations. 4.8 Factors relating to parents and carers of young people who experience running away and CSE Professionals identified the following issues: - The difference that supportive parents and carers can make for young people, and therefore the importance of being able to work with parents and carers, including stressing the importance of reporting young people who run away as missing to the police. - There is a lack of resources to work with parents and carers of young people experiencing running away and CSE.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		- Professionals need to ensure that they do not seem to be 'blaming' parents and carers for the young people's runaway behaviour. 4.9 Factors relating to direct practice with young people who experience running away and CSE 4.9.1 Factors supporting direct practice The study reports that: Young people identified that: - There is a need for more services 'where they can just turn up' (p82). One young person said: 'There should be places where kids can go to tell someone what's happening to them – someone who will believe them and be able to help them and know what to do for the best. I know there's ChildLine and that but there's some things you don't want to say over the phone. Kids want to go somewhere where there's people they can talk to face-to-face' (p82). - It can be easier to trust workers from the voluntary sector than those from statutory agencies. - It's important for professionals to listen to them. - Outreach work can be valuable. - A good relationship with the worker is 'paramount' (p85), this is supported by an informal manner, and professionals doing what they say they will do. - That it is not always possible to stop young people from running away, but that support should continue to be provided. One young person said: 'I think you have to keep that support in place even when the young person isn't listening and continues to run away and have sex with older men so that when they	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		realise what is going on, there is support in place for the young person … Don't just brush them aside because they can't be bothered at the time. To be honest, it's at the time that they can't be bothered with the support when they don't realise the situation that they're in' (p87). - Ensure a focus on the future as well as the past. - Where they had experience of peer support they had found this helpful. One young person said: 'I wasn't taking any notice of what X [the young person's support worker from the specialist sexual exploitation project] was saying so he got this guy who was 18 and had similar experiences as me to come and talk to me … It helped because it was like another view of what I would see in three years' time … He [the 18-year-old male] was like "I know how it is: it's like the best thing in the world and you think they perpetrators] all love you but they do not; they genuinely do not love you; they don't care about you". And it did help me because he had been through this.' (pp88–9). - Young people liked to express themselves using creative outlets. Professionals identified that: - Young people should be able to self-refer to services, and that services should be provided in a 'warm and friendly' environment (p84). - Outreach work can be valuable. - Young people appreciate long term involvement from a consistent worker. - Taking time to build relationship and engage with young people is important. - Flexibility of approach is important, to suit the needs of different young people.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		4.9.2 Factors that hinder direct practice Young people identified that: - They would like to see their social worker more often, and felt that statutory practitioners did not always communicate and keep them informed. Professionals identified that: - It is unhelpful to have to stop working with young people when they turn 18. - Some young people are particularly difficult to engage, including: those who are not reported as missing; those who run away for long periods of time; young people who have 'become entrenched in red district culture' (p90); Roma children and young people; 16–18 year olds who have no statutory involvement; young people who have 'fended for themselves' for some time and have become mistrustful of agencies; heterosexual males who are unwilling to disclose exploitation; younger children, such as those aged under 11; young people who have had a lot of professionals involved with them; those who have not developed trusting relationships with any adults; young people with a late diagnosis of ADHD. - Direct work can be hindered by young people's own lack of recognition that they are being sexually exploited. One professional said: 'One of the most difficult ones [hindering factors] is that young people don't recognise their exploitation and so that is a challenge in itself … The young people are needy, they want love and a sense of belonging and that's what exploiters home in on … It may be the only love they've [young people] experienced is in the context of sexual	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		abuse. Some of these children don't have any experiences of safe supportive adults and they distrust professionals, obviously. It's not a quick fix: some of them have really poor attachment history; there needs to be long-term work' (p92).	

20. Stalker K, Green Lister P, Lerpiniere J et al. (2010) Child protection and the needs and rights of disabled children and young people: a scoping study: abridged report. Glasgow: University of Strathclyde

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The objectives were to: - scope current knowledge about child protection and disabled children - review current social policy and practice in the field, and - pilot ways to seek disabled children's views about the child protection system (p6). **Methodology** Qualitative study. Relevant section is from 10 interviews with key informants. Purposive sampling was used to select key informants expected to have close knowledge of policy issues relating to child protection and disability. These interviews	**Participants** Professionals/practitioners - 10 interviews with key informants selected for their insight into child protection and disability. The representatives were from: senior policy makers and practitioners based in central government, the inspectorates, the police, the NHS, the voluntary sector and a Children's Commissioner Office. **Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported.	**Narrative findings** Factors that help and hinder professionals' effective response to meeting the needs of children with disabilities: 1. Communicating with disabled children: - Many participants identified that communication was a challenge, but was the professionals' responsibility to engage with child. As 1 participant commented: 'All children can communicate something and [professionals] shouldn't ever dismiss the possibility of getting information from children if you find the right way' (p17). - An inspector commented on social workers who direct questioning to parents, rather than the child. This was seen as a lack of confidence, knowledge and experience. 'The extra time needed to interview many disabled children could be problematic where investigations had to move quickly or specialist support workers were not readily available, especially in rural areas' (p17). - Good examples of practice were presented in the form of a case example, where a professional was cited as doing 'some very skilled work ... with the	**Overall assessment of internal validity:** - No information on data analysis, collection or how findings are contextualised. Does provide an overview of what is currently known and meets research aims. **Overall assessment of external validity:** + Meets criteria. The participants generally represent Scotland (n=8) so caution to generalise the UK. **Overall validity rating:** - Poor research design

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
were carried out with senior policy makers and practitioners in Scotland (n=8) and England (n=2). Unfortunately the research team could not interview anyone from Wales or Ireland. **Country:** UK, Scotland (n=8). England (n=2). **Source of funding:** Not reported.	• Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** n=10.	child, using games, toys and other methods to help her explain what had happened and as a result she started talking' (p18). 2. Under-reporting of abuse: - One participant spoke with social services managers and it was on this basis, commented that social workers tend to focus on the relationship parents, than the child, as they see the parents subsequent stress of caring and are disinclined to register formal child protection concerns because it is 'a wee bit of neglect and whatever' (p18). Coupled with an inspector who commented on poor practice which led her to believe some social workers are more tolerant of parents smacking disabled children than other children. - Three participants from inspectorates raised the concern of the paucity of information about disabled children in child protection inspections and reports. 3. Differential treatment within the child protection system: - Several participants stated that disabled children were often poorly served. A case example is cited by one key informant from her practice experience where 'a young boy was abused by two adults in his home ... The boy's aunt found out about it and removed him from the house. A joint child protection investigation began and the boy was very clear about what happened ... The investigating team decided to take no further action, despite the boy having given full description of what had happened' (p19). 4. Joint working: - It was considered good practice to have effective joint working measures in place, and due to children	with little information about how data is collected, analysed and thus, conclusions are difficult to draw.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		with disabilities having additional support from other agencies, participants commented shared responsibility in safeguarding had improved over the years. - Child protection teams and child disability teams were reported to lack knowledge of each other and that impairment was not always adequately reported or consistent. - Resources and ring-fencing funds for protecting children with disabilities impact on provision which was seen as unevenly available across the country.	

21. Stanley N, Miller P, Richardson Foster H (2012) Engaging with children's and parents' perspectives on domestic violence. Child and Family Social Work 17: 192–201

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The study states that: 'This research examined both the notification process itself and the subsequent service pathways followed by families brought to the attention of children's social services in this way. It also explored which other agencies contributed to services for families experiencing domestic violence and captured young people, survivors' and perpetrators' views of services' (p9). For the	**Participants:** Children and young people - 5 focus groups with children and young people aged over 10 (precise number not reported in this paper, however Stanley et al. 2012 report that 19 young people were involved in total). Caregivers and families - Interviews with adult survivors and perpetrators. Numbers not reported in this paper. However, Stanley et al. 2012 report that 11 domestic abuse survivors were involved, and 10 perpetrators. **Sample characteristics:** • Age - Children and young people: Ages ranged from 10 to 19 (reported	**Narrative findings** Section 3.3 Experiences and perceptions of police intervention in domestic violence 3.3.1 Attitudes towards the police The study reports that young people, survivors and perpetrators showed a range of attitudes towards the police, with many feeling distrustful. Survivors felt that good police officers had good communication skills and were able to 'show empathy and sympathy' (p49). BME participants reported different experiences. Some felt that their race meant that they had not received as good a response, whereas others felt that they had been treated respectfully. 3.2.2 Speed of police response	**Overall assessment of internal validity:** - **Overall assessment of external validity:** + **Overall validity rating:** - Very little information given regarding participants' demographic characteristics. Little description or justification of sampling. Unclear why focus groups

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
purpose of our review, the second part of the research question, relating to the experiences of service users, was more relevant. **Methodology:** Qualitative study. Qualitative study comprising five focus groups with children and young people and interviews with adult survivors and perpetrators of domestic violence. Where relevant, additional methodological detail has been taken from Stanley N, Miller P and Richardson Foster H (2012) Engaging with children's and parents' perspectives on domestic violence. Child & Family Social Work, 17: 192–201. **Country:** UK, England.	in Stanley et al. 2012) Survivors: Aged between 25 and 48 years, with a mean age of 38 years (reported in Stanley et al. 2012) Perpetrators: Aged between 30 and 45 (reported in Stanley et al. 2012). • Sex - Children and young people: 11 female, 8 male (reported in Stanley et al. 2012) Survivors: 10 female, 1 male (reported in Stanley et al. 2012) Perpetrators: 0 female, 10 male (reported in Stanley et al. 2012) • Ethnicity - Children and young people: 16 White British, 1 White Asian, 1 White/Black Caribbean, 1 White/Black African. Survivors: 4 White British, 7 described themselves as black or minority ethnic (BME) groups (reported in Stanley et al. 2012) Perpetrators: 6 White British, 4 BME groups (reported in Stanley et al. 2012). • Religion/belief - Children and young people: Not reported. Survivors: Not reported. Perpetrators: Not reported. • Disability - Children and young people: Not reported. Survivors: Not reported. Perpetrators: Not reported. • Long term health condition - Children and young people: Not reported. Survivors: Not reported. Perpetrators: Not reported.	Young people, survivors and perpetrators all thought that the police were sometimes slow to respond to cases of domestic violence. 3.3.3 Provision of information and explanations Young people in particular thought that they did not get sufficient information and explanations from the police. One young person said: 'When my dad came round and he started kicking off, the usual after a while, the police come round and they arrested him, they took a statement off my mum and that's it, they don't … they didn't say to us what happened if he was going to be released the next day or we didn't find out anything' (Dawn, young people's focus group 4, p52). Survivors also thought it was important that they were provided with sufficient information. Perpetrators also appreciated clear information about what would happen next. 3.4 Being listened to and validated Young people, survivors and perpetrators all reported that they did not feel sufficiently listened to by the police. 3.4.1 What happens next? A number of survivors reported that they had not been satisfied with the follow-up actions following a police response to a domestic violence incident. Several young people thought that the perpetrator should be removed from the scene of the incident as soon as possible. One young person said: 'When they come straight away, they could, like, take him away	used for some participants but interviews used for others.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	• Sexual orientation - Children and young people: Not reported. Survivors: Not reported. Perpetrators: Not reported. • Socioeconomic position - Children and young people: Not reported. Survivors: Not reported. Perpetrators: Not reported. • Type of abuse - Children and young people: Not reported. Survivors: Not reported. Perpetrators: Not reported. • Looked after or adopted status - • Unaccompanied asylum seeking, refugee or trafficked children - Children and young people: Not reported. Survivors: Not reported. Perpetrators: Not reported. **Sample size:** Authors state that a total of 40 young people, survivors and perpetrators were involved. A breakdown for each group is not given.	straight away, instead of waiting around and everything and listening to sides, just … they should be taken away because a mum or child wouldn't call 999 just to get a dad taken away for no reason' (Louis, young people's focus group 5, p58). 3.4.2 Continuity of support All 3 participant groups emphasised the importance of continuity of support following an incident of domestic violence. 3.4.3 Barriers to accessing help and support from the police Survivors highlighted the following barriers: - Intimidating environment of the police station. - Confidentiality not always protected [assume this also relates to at the police station]. 3.5 Experiences and perceptions of children's social care 3.5.1 Children and young people's experiences and perceptions of social workers Young people reported a range of perceptions of social workers, from very positive to very negative, but particularly appreciated those who made themselves available and talked directly to young people. Negative experiences tended to relate to social workers who were perceived not to listen to young people and their families. Some young people thought that social workers put pressure on them to give information. One young person said: '... they try and get it out of you ... they keep pressuring you to get answers' (Richard, young people's focus group 2, p61).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Some young people spoke negatively about social workers who had not 'kept their promises' (p61). 3.5.2 Survivors' experiences and perceptions of social workers The study reports that many of the survivors involved had had little contact with social workers, and some had had no contact. Many survivors expressed fears about their children being taken away, including 1 woman who had significant experience of children's social care services who said: '… there was a time when I called them, when there was violence carried on, and they didn't do anything like [they would now]. "The only thing that we can do if you continue to allow this man into your home, and he continues to be violent towards you, that we will have to take your little boy away'. That's what they said, they will take children away"' (Rose, survivor, p62). Some participants said they were confused about what the role of social care services were, and whether they were able to help them to access support. 3.5.3 Perpetrators' experiences and perceptions of social workers Most perpetrators had little experiences of social workers. Of those who did, some said that they felt they had not received good communication from the social worker, but had had to hear messages second hand from their ex-partner. A small number of perpetrators thought that the social workers had communicated well with them, and also that child protection conferences and family group conferences had been helpful.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		3.6 Experiences and perceptions of specialist domestic violence services 3.6.1 Young people's services The young people taking part in the study had been recruited via specialist domestic violence services for young people. Young people were very positive about these services, and the fact that the workers listened to them, and took the time to build up relationships with them. 3.6.2 Refuge services The study reports that over half the survivors interviewed had used at least one refuge, and reported valuing 'feeling safe'. Survivors reported that refuges varied in the extent to which they supported children, and thought more could have been offered in terms of counselling or interventions for children and young people. 3.6.4 Community domestic violence outreach services The study reports that survivors valued the practical resources provided through community outreach services, as well as outreach counselling or helpline services. 3.6.5 A multi-agency domestic violence service Survivors in one site had access to a multi-agency service offering legal, health and advice services. This was seen as valuable. 3.6.6 Perpetrator programmes Out of scope. 3.7 Experiences and perceptions of other services	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		3.7.1 Probation services The study reports that 6 of the 10 perpetrators had had contact with probation service, and valued the 'close, non-judgemental relationship' (p69) this provided, as well as flexibility with regard to appointment times. 3.7.2 Legal services and courts The study reports that the survivors interviewed had general had positive experiences of solicitors on the court system, and valued specific advice from solicitors on how to keep their children safe (e.g., writing to the school to ensure that their ex-partner did not collect their child). One survivor reported negative experiences of the court system. 3.7.3 Health services Survivors who had disclosed domestic violence to their GPs had generally found them helpful and supportive. Health visitors were also named as a useful source of support. 3.7.4 Education Young people had mixed experiences of getting help from teachers in the context of domestic violence. Two young people said: 'No, teachers don't have the time for you. Because it's not their job, and just … And they're not prepared' (Nicola and Tupac, young people's focus group 1, p71). Other young people thought that the response offered by other education professionals was inadequate. One young person said: 'And the school counsellor was exactly the same as that, she said, she asked	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		what was happening, and when you tell her she just sits going, “um, yeah”’ (Tanya, young people’s focus group 2, p71). The study reports that, generally, there was little mention of schools is survivors’ or perpetrators’ descriptions of the service response, except one survivor whose child had received extensive support from the school. 3.7.5 Counselling/therapy services Some children had been offered therapy or counselling services, which their parents (survivors) thought had been helpful. A number of survivors had accessed therapy themselves and had found this valuable.	

22. Taylor J, Stalker K, Fry D et al. (2014) Disabled children and child protection in Scotland: investigation into the relationship between professional practice, child protection and disability. Edinburgh: Scottish Government Social Research

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The study reports that ‘the aim of this study was to assess how public services (including social work, health care, education, police and other related services) identify and support disabled children and young people at risk of significant harm,	**Participants:** Professionals/practitioners - Practitioners ‘working on issues of disabled children and child protection’ (p.1). Data was gathered from 21 practitioners across six local authorities via interview, and the remaining 40 through focus groups with five local authority Child Protection Committees. The roles of the practitioners involved are not clear. The research re-	**Narrative findings** Data have been extracted from Section 5 ‘The child at the centre’ and Section 8.3 ‘Are we getting it right for every child?’ 5.1 Child centredness The study reports that participants emphasised that ‘every child, whether disabled or not, should be seen first as a child, thereafter as a child with an impair-	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Limited information on ethical considerations in relation to focus groups.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
whether neglect or abuse' (p13). The study had the following four research questions: '1. What are the decision-making process and 'triggers' for intervention used by professionals when determining the nature of interventions for disabled children and young people at risk of significant harm? 2. What are specific issues faced by practitioners in Scotland in supporting children and young people at risk of significant harm? 3. How do services co-ordinate to support disabled children and young people at risk of significant harm? 4. What are practice examples in Scotland addressing these issues?' (p13). The findings are not structured according to the 4 research questions, but according to 3 over-arching themes which are: 1) The child	port states that 'from each local authority area, potential participants were contacted from social work, education, police, voluntary organisations and health with practice experience of responding to at least two child protection cases involving a disabled child' (p. 14). However it is unclear who was actually recruited to the study. **Sample characteristics:** • Age - Not reported. • Sex – Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** n=61.	ment' (p20). Participants distinguished between disabled children with communication impairments, for whom it was thought that behaviours signalling harm would be different from the general population, and those without communication impairments, who were expected to make a 'disclosure of abuse' (p20). The authors note that this suggests a misunderstanding of disability, and the impact that it can have on children other than just their ability to communicate. The study notes that participants stressed the importance that disabled children at risk of or experiencing significant harm were included within the child protection system framework. Participants also noted the importance of taking an individualised approach to each child. 5.2 Impairment effects The study reports that there was a distinction between participants who thought that disabled children faced unique risks, and those who thought they would take the same approach as for any child. Participants did not acknowledge that the impairments experienced by disabled children could add complexity to a child protection case. One participant said: 'I think our rate of detection is probably quite poor because I think of all the personal care and things that child have, I would suspect that the rate of sexual abuse and stuff is probably higher than we actually detect. It's hard enough in the average population without them being disabled where they can't talk and tell us [Interview 3]' (p22). Participants highlighted that, in some cases, it was difficult to assess risk because it was unclear whether	**Overall validity rating:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
at the centre; 2) Practice issues (muddling through) 3) Inter-agency working. We have data extracted findings in relation to theme 1 only, as both 2 and 3 relate more to organisational configuration and training issues. **Methodology:** Mixed methods. Research comprised 'in-depth' interviews with 21 practitioners which included use of a Critical Incident Technique methodology, and five focus groups with Child Protection Committees. From the interviews with practitioners, 34 practice examples were developed. The team also developed a series of models to represent the data from the interviews and focus groups. **Country:** UK, Scotland.		a behaviour was due to a (potentially undiagnosed) disability or due to a child protection issue. 5.3 Communicating with disabled children The study reports that practitioners appeared to associate recognition of abuse primarily with disclosures by children. The authors note that 'Waiting for a disclosure is a reactive stance to child protection' (p25). The study notes that, although participants thought that disabled children with communication impairments would be less able to disclose abuse, in fact in a number of the case examples disclosures had been made by children with communication impairments. The study notes that participants thought that 'due to a lack of knowledge and training, or a perceived inability of children to communicate, there was a greater chance of missing signs of neglect and abuse that would be picked up more efficiently in non-disabled children' (p26). The study notes that it is concerning that some interviewees appeared to imply that this was due to the child's lack of ability, rather than the professional's. Despite this, a number of examples were given of adapting communication, including use of speech and language therapists. 5.4 Child agency The study notes that, even in instances were communication was adapted, a child protection case did not always progress as it would for non-disabled children. 5.5 Parents	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Source of funding: Government - Scottish Government.		The study notes that participant's thought that the presence of a disabled child affected decisions about whether situations were neglectful, or whether parents required more support. Participants also noted that parents' desire to protect their children could also lead to situations of unintentional neglect. Participants thought that parents may find it hard to admit they need more help if the situation changes, or may have negative perceptions of social services. 5.6 The invisible child The study notes that participants expressed concerned that child protection practitioners may occasionally over-empathise with the parents of disabled children, and thereby underestimate the risks posed to the child. Participants said that there were often a wider range of services involved with disabled children. This could have the positive effect of ensuring that abuse/neglect didn't go unnoticed, but could also lead to complacency that someone else may be dealing with it. Participants also noted that parents and carers sometimes had to be used as a proxy for communicating with a disabled child. The authors give the following summary of themes: - Participants interpreted current policy as meaning that disabled children thought to be at risk should not be treated differently from other children. The authors note that this 'did not always translate into effective identification and intervention for child protection risks involving disabled children' (p36).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		- The language of treating every child the same did not match the reality of some of the practice issues described - and it was clear that adaptations were (rightly) made for disabled children. - Participants identified adaptation of buildings and service provision as important for effective protection of disabled children. Participants differed in the extent to which they thought that services provided a good response to disabled children. However, there was more consensus that there was a lack of suitably adapted service provision, for example residential accommodation and placements. 8.3 Enablers and barriers: Are we getting it right for every child? Note - this section appears in the Discussion section of the report. The study identifies the following enablers and barriers to working with disabled children throughout the child protection process (information taken from Figure 6. p71): 1. Concerns raised Enablers - Interagency working, passing on concerns Barriers - Complex family situations, parent support/carers around child, communication impairments, lack of knowledge on impairments and support needed, fear of 'getting wrong'. 2. Initial information-gathering	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Enablers - Initial Referral Discussion process, inter-agency working, good information sharing, working with specialists Barriers - Difficulty accessing interpreters and other support, lack of CYP involvement, overreliance on third party information, information held across teams/areas resulting in delays, lack of worker confidence and knowledge. 3. Decision to launch investigation Enablers - Interagency working Barriers - Burden of evidence needed, communication impairments, CYP deemed as 'unreliable witness'. 4. Planning Enablers - Interagency working Barriers - Lack of accessible support (interpreters, speech and language therapists etc.); 'muddling through' case reflecting lack of clarity, knowledge and experience; CYP passed on to other services. 5. Child Protection Case Conference Enablers - Interagency working Barriers - Support for CYP involvement, inaccessible venue spaces, lack of knowledge around types of impairment. 6. Child Protection Plan Enablers - Interagency working Barriers - Child protection concerns for other CYP not address, burden of proof (specifically in relation to communication impairments), difficulty in finding foster carers for disabled children.	

23. Wirtz L (2009) Hidden children: separated children at risk. London: Children's Society

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To explore the experience of trafficking and abuse of migrant young people in terms of disclosure, intervention and their support needs. **Methodology:** Qualitative study. Interviews and focus groups with trafficked migrant children, professionals and analysis of 34 case studies. **Country**: UK. **Source of funding:** Voluntary/charity - City Parochial Foundation and East Foundation.	**Participants** Children and young people. Trafficked migrant children and young people Professionals/practitioners. Professionals and practitioners working with vulnerable young people from the voluntary and statutory sectors. **Sample characteristics** • Age - Trafficked migrant children: aged 16 years and over. Professional practitioners and volunteers: not reported. • Sex - Trafficked migrant children: 40% boys and 60% girls. Professional practitioners and volunteers: not reported. • Ethnicity - Trafficked migrant children: from 4 Asian and 10 African countries. Professional practitioners and volunteers: not reported. • Religion/belief - Trafficked migrant children: not reported. Professional practitioners and volunteers: not reported. • Disability - Trafficked migrant children: not reported. Professional practitioners and volunteers: not reported. • Long term health condition - Trafficked migrant children: not reported.	**Narrative findings** A. Disclosures: 1. Unintentional disclosures - when trafficked children approached practitioners for help and advice on issues (such as housing, immigration) and the story came out during the discussion, or the disclosure may follow general enquiries or small talk such as 'how are things at home' (p37). 2. Wanting to disclose but waiting until they are safe and with a person they trust, suggesting the importance of relationships of trust enabling disclosures, suggesting that trafficked children might be more open with volunteer agencies and social workers than immigration officials. 3. Disclosure taking place when young people feel their situation cannot get any worse such as when they were sleeping on the street, when they were pregnant and afraid that their baby would also be abused or after an incident when they had been particularly badly physically abused (p37). 4. Knowing about other people's experiences could empower young people to seek help themselves, such as when a friend disclosed her own abuse, this gave them the courage to say 'me too' (p39). 5. Young people's contact with their families back home was considered as a possible means of disclosure and escaping their situation.	**Overall assessment of internal validity:** ++ **Overall assessment of external validity:** ++ **Overall validity rating:** ++

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Professional practitioners and volunteers: not reported. • Sexual orientation - Trafficked migrant children: not reported. Professional practitioners and volunteers: not reported. • Socioeconomic position - Trafficked migrant children: 'The children came from capital cities as well as villages with no school and no electricity' and 'from countries affected by war or political violence' suggesting unstable social economic background. Professional practitioners and volunteers: not reported. • Type of abuse - Trafficked migrant children. Professional practitioners and volunteers: not relevant. • Looked after or adopted status - Trafficked migrant children: not reported. Professional practitioners and volunteers: not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Trafficked migrant children. Professional practitioners and volunteers: not relevant. **Sample size** Trafficked migrant children: n=8. Professional practitioners and volunteers: n=15.	6. The young people felt that the following conditions might have got them out of the situation sooner - If trusted adults asked direct questions (yes/no questions) if they were being abused; if people in the community hadn't been afraid of their guardians; if people weren't afraid they would get a criminal record for helping; if something had made them frightened enough to run away sooner; if a service who knew what their rights were had known about their home situation; if social services had rejected their guardian as a private foster carer because of a history of mental illness (pp39–40). 7. Practitioners' felt that the following conditions could help trafficked children to disclose sooner- Having access to services (knowing help is available) and something could be done about their situation; if they had known trafficking was wrong; if they had information at school; if questions had been asked by the school about who they stayed with and what the relationship was; if they knew they wouldn't be made homeless, deported or put in prison by the immigration authorities; if a properly resourced private fostering team had looked at all the evidence; if a support plan had been done earlier; if they knew someone they could trust; if they knew they weren't the only children being trafficked; if the school had informed social services that they were privately fostered; if children's services had actively investigated the case; if they believed children's services will protect them; and if social services had followed up or monitored the private fostering arrangement and spoken separately with the young person (p40). B. Locations for interventions (p40)	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Depending on how well informed trafficked children were and the people they met there, they usually visited these places when still under the control of their exploiters and/or immediately after escaping - Immigration service on arrival at UK airports; Children's Services; churches; Connexions; dentists; ethnic shops e.g. African or Vietnamese supermarkets; GP surgeries; hospitals; job centres; migrant community groups; youth advice centres and youth clubs; nurseries and play centres; police stations; school/college; solicitors. This showed the range of people and agencies that could be the first point of contact or disclosure for hidden children, suggesting the importance of a range of agencies having access to training on trafficking. 1. Possible locations for interventions - most likely to be in informal community (migrant community groups) and faith settings (such as churches) in the community. 2. Interventions could be made more likely by increasing public awareness in the community of children's rights and UK law regarding child safeguarding and private fostering, the Community Partnership Project (CPP) being a good example. Also through television advertisement, the Outreach through training and advice, community meeting (involving the police, social workers and local residents), public awareness campaigns (such as the Blue Blindfold Campaign; the Poppy Project). 3. Interventions in school - secondary school settings with school counsellors and advisors an important resource. Practical support outcomes, with referrals to	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		other services were felt to be more important than having someone to talk to. Confidentiality of these services should also be made explicit (p44). 4. Interventions in healthcare settings - trafficked children should not face barriers when trying to access healthcare, e.g., denying medical treatment such as immunisations puts the health of the wider community at risk (p45). 5. Interventions by immigration staff - new resources needed toward improving the detection/identification of trafficked children at ports and airports. C. Support needs - practitioners reported an ongoing misunderstanding of entitlements for migrants and that social workers assume any foreign young person has no recourse to public funds. Hidden children were often caught in disputes between two local authorities that each said the other was responsible for providing care (p46). 1. Housing - trafficked children taken into local authority care or into new private fostering arrangements and outcomes had not been always successful. 2. Therapeutic support - hidden children likely to perceive counselling only involved talking which might not offer those practical help. Some western talking therapies might be culturally inappropriate and a culturally sensitive therapeutic casework service could be of benefit (p47). 3. Post-16 support - hidden children usually have no family in this country, so after age 18 or 21 they may have no one for practical or emotional support. When they turn 18, unless they are NEET (not in employment, education, or training) their support sometimes stops suddenly. Without this support, some hidden	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		children unwittingly became unlawfully in the country (p48). 4. Preparation for parenthood- some practitioners felt that hidden children may have lacked positive parental role models if they spent much of their lives in situations of exploitation and abuse, and so care has to be taken to ensure they are equipped to give their own children the care that they themselves did not receive, especially when they get pregnant at an early age, often soon after escaping domestic servitude. To prevent future exploitation, there might be a need for young people to be aware of the law, the importance of their own consent, and what constitutes a healthy relationship (p49). 5. Some young people who had become mothers cited their children as a positive aspect of their current lives, including them in their protection shields as something that makes them feel happy and strong (p49).	

Organisational factors

Review question 21 – What organisational factors support and hinder effective multi-agency working including supporting good professional judgement?

Review question 21 – Critical appraisal tables

1. **Beckett H, Brodie I, Factor F et al. (2013) 'It's wrong … but you get used to it' - A qualitative study of gang-associated sexual violence towards, and exploitation of, young people in England. London: Office of the Children's Commissioner for England**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. The research	**Is the context clearly described?** Clear. Under each direct quote, it	**Does the study's research question match the review question?**	**Overall assessment of internal validity:** ++

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
team conducts individual interviews with young people (n=150); 11 focus groups with professionals (n=76); and 8 single-sex focus groups (n=38). The comprehensive methods section details the rationale for interviewing participants because of the sensitive nature of the topic and to follow an ethical protocol. In addition, safeguarding concerns have been explored. **Is the study clear in what it seeks to do?** Clear. The forward from Sue Berelowitz, Chief Executive, Office of the Children's Commissioner details the context of the research: very little is known about the prevalence of sexual violence and exploitation within gangs by children and young people against other children and young people. The purpose is to understand through interviews with young people and professionals' experiences to better inform national and local policy. **How defensible/rigorous is the research design/methodology?** Defensible. Very thorough research design and methodology which was governed and reviewed by a number of different bodies:	is clear where data was collected, whether they are a young person or professional and age of participant (if individual interview). The individual interviews with young people (n=150) contain detailed characteristics, however the focus groups held with professionals (n=74) and young people (n=38), it is unclear on the characteristics of these participants. **Was the sampling carried out in an appropriate way?** Appropriate. Very clear that participants were recruited via agencies that were supporting young people to minimise risk. The authors state the potential for 'bias into the sample - and excludes other potential participants with valid contributions to offer - it was felt that the risks of engaging those outside of services could not be adequately negated within a time-limited, large-scale, multi-site project such as this' (p12). **Were the methods reliable?** Somewhat reliable. The data is collected by one method, which were qualitative interviews.	Yes. The study explores 150 young people's and 76 professional's responses to gang-associated sexual violence and exploitation. The purpose is to understand the prevalence and experiences of young people: Chapter 4 is relevant to research question because it explores preventative strategies, including the impact of multi-agency working. **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval was gained from four different Research Ethics Committees and relevant local approvals were obtained within each research site. The research team was accountable to a Research Project Advisory Group, a Young People's Advisory Group and local Multi-agency Advisory Groups in each research site. **Were service users involved in the study?** Yes. In order to use age-appropriate research questions, the Young People's Advisory Group co-produced the interview schedule. **Is there a clear focus on the guideline topic?** Yes. The relevant section is Chapter 4.2 where	Very comprehensive, effective study with detailed findings that are summarised into recommendations. The methodology has dealt appropriately with the ethics and risk associated with the subject matter, and the research team have ensured the voice of the child is at the heart of the report. **Overall assessment of external validity:** ++ The study meets all criteria and has dealt effectively with ethical considerations. In addition, the interview guide was co-produced with the YPAG to make questions age appropriate. **Overall validity rating:** ++ A thorough empirical study which meets its research aim and details implications for practice and policy on a local and national level.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Research Project Advisory Group; Young People's Advisory Group' and Site specific Multi-agency Advisory Groups. Qualitative interviews were conducted with 150 young people; 11 focus groups with 76 professionals; and 8 single-sex focus groups with 38 young people. There is a detailed breakdown of the 150 young people who participated in individual interviews, however the focus group held with professionals and young people is not descriptive. **How well was the data collection carried out?** Appropriately. Data collection section is thorough and the research team explained the measures to ensure the participants comfortability by facilitating the young people to talk in the third person, unless they wanted to actively choose otherwise, i.e. conversational manner using the interview schedule as a framework for discussion. There is effective consideration of the commitment to maintaining participants' confidentiality and anonymity. An ethical protocol was developed on the basis of 'no harm should come to any individual as a result of their agreement to facilitate or take part in the work' (p12).	**Are the data 'rich'?** Rich. The research team cite references to where each finding was collected which helps contextualise responses to each participant. There are limitations as explored: 'Due to the flexibility built into the interviewing process, not all issues were covered with all of these interviewees' (p14). **Is the analysis reliable?** Reliable. Qualitative interviews were thematically analysed using NVivo 8 which underpin the findings in the research. The research team explain the executive decision to generally prioritise the young persons' voice to be presented in the report. **Are the conclusions adequate?** Adequate. The narrative findings of the voice and experience of participants contextualise the current knowledge and prevalence of sexual violence and exploitation in gangs. The relevant section to disclosure (Chapter 4) concludes that from the aim of 'identifying learning for embedding more effective systematic response to these issues in the future ... Prompted re-	young people and professionals state factors that hinder disclosure: confusion about what actually constitutes sexual violence and exploitation; the acceptance of sexual violence and exploitation; and low levels of reporting and seeking support from professionals, i.e. judgement by others, lack of faith in services, perception of police and absence of conviction. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. 150 Young people's experience of gang-associated sexual violence and exploitation, and professionals (n=76) who have experience/specialism working with sexual violence and exploitation. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Young people were selected because they were/had received support from services, and professionals from statutory services were interviewed, i.e. social care, police, and education. **Does the study relate to at least 1 of the activities covered by**	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	sponses to these are now presented in the form of recommendations' (p51). The recommendations are structured to address national and local policy, which in the context of presenting findings from 6 different localities in England, map the issue with scope to respond.	**the guideline?** Yes. Chapter 4 relates to barriers to professionals for young people disclosing sexual violence and exploitation. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. 'The research aimed to consider: the scale and nature of gang-associated sexual violence and exploitation in six areas of England; the main pathways into gang-related sexual violence and exploitation for young people living in these neighbourhoods; and potential models for an effective multi-agency response to the issue' (p6). **Does the study have a UK perspective?** Yes. Study carried out in 6 different research areas. For confidentiality purposes the sites are not named but do 'reflect a broad range of experiences of working with gangs and different demographic profiles' (p6).	

2. **Berelowitz S, Clifton J, Firmin C et al. (2013) 'If only someone had listened': Office of the Children's Commissioner's inquiry into child sexual exploitation in gangs and groups. London: Office of the Children's Commissioner for England**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Defensible. **How well was the data collection carried out?** Appropriately. Data collection: 1. Call for evidence from various child protection agencies dealing with CSE 2. Dataset requests to gather evidence of the extent and nature of practice which has been developed by LSCBs, police forces and specific health services (GUM Clinics, Substance Misuse and CAMHS) 3. Semi-structured interviews and focus groups with children, young people, parents and carers 4. 10 site visits 5. Workshops and academic seminars.	**Is the context clearly described?** Clear. **Was the sampling carried out in an appropriate way?** Not sure. Not reported. **Were the methods reliable?** Reliable. Multiple methods of data collection. **Are the data 'rich'?** Rich. **Is the analysis reliable?** Reliable. Relating to qualitative data collection 1. The interviews, site visits, evaluations, workshops and seminars were analysed thematically and used to identify trends. Interviews recorded and content analysis, using NVivo 10. 2. Qualitative data analysed manually and in the case of questions asked around barriers, the same coding framework was used across the dataset and the call for evidence to enable key consistencies/variations across the evidence captures to be identified.	**Does the study's research question match the review question?** Yes. To assess views and experiences of victims of and professionals working in CSE 1. Why CSE children continue to be let down 2. Why there is a no fully joined-up multi-agency, child-centred approach to address child sexual exploitation in gangs and groups, why agencies and individuals fail to listen to them, and fulfil their responsibilities with regard to child protection, or that there was not sufficient strategic and managerial oversight to coordinate their actions. **Has the study dealt appropriately with any ethical concerns?** Partly. Insufficient information provided. 'Letters of thanks were sent to the professionals and any participating young people, and any specific ethical concerns or safeguarding issues were followed up' (p120). 'Interviews were recorded subject to approval' (p120).	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + Overall methodology of this study was sound though some details missing (sampling and population characteristics). Data analysis was robust and interpretations of findings convincing.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	Are the findings convincing? Somewhat convincing. Similar to the analysis, without a clear framework identified and the authors stating that they could not 'include all the detailed information' (p26), it is difficult to form a base of judgement on whether the findings are reliable. **Are the conclusions adequate?** Somewhat adequate. Again, the conclusions highlight overarching themes but there is no consideration of limitations or clarity where data is from, i.e. police officer or social worker, hence making conclusions difficult or reliable.	**Were service users involved in the study?** No. Service users did not co-produce this report **Is there a clear focus on the guideline topic?** Yes. Child sexual exploitation (CSE). **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. CSE children. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. CSE. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes. UK.	

3. **Brandon M, Sidebotham P, Bailey S et al. (2013) New learning from serious case reviews: a 2-year report for 2009 to 2011. London: Department of Education**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Defensible. Nature of serious review: qualitative overview reports and/or executive summaries of 5 studies. **How well was the data collection carried out?** Appropriately. Overview reports and/or executive summaries of 5 individual but interlinking studies to provide a thematic and critical analysis of recommendations and action plans from 30 serious case reviews.	**Is the context clearly described?** Clear. SCR. **Was the sampling carried out in an appropriate way?** Not sure. **Were the methods reliable?** Somewhat reliable. **Are the data 'rich'?** Rich. **Is the analysis reliable?** Reliable. A triple-layered reading process was carried out for each case. 1. Redacted overview report was read and summarised to identify key points and produce a structured summary sheet 2. The overview report was read again and data were coded to the most appropriate nodes within a thematic coding framework 3. Data were analysed thematically within three core domains: the child; family and environment, including parenting capacity; systemic and service issues. Three researchers were involved in reading, coding and analysis, and the team used a constant comparative approach to	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** No. Not reported. **Were service users involved in the study?** No. Service users did not co-produce this report. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children with serious and fatal maltreatment. **Is the study setting the same as at least 1 of the settings covered by the guideline?** No. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Child maltreatment. **(For views questions) Are the views and experiences reported**	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** + This serious case review analysed national data based on a large sample of cases, with fairly sound methodology.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	analysing data, looking for emerging themes and outliers. Team discussions were held to identify common emerging themes. Software package NVivo9 used for analysis. **Are the findings convincing?** Convincing. **Are the conclusions adequate?** Adequate.	**relevant to the guideline?** Yes. Serious case reviews with data gathered from 5 studies, to explore interacting risks which seeks to understand inter-agency working within the dynamic context of the developing child's world. **Does the study have a UK perspective?** Yes.	

4. **Brodie I and Pearce J (2012) Exploring the scale and nature of child sexual exploitation in Scotland. Edinburgh: Scottish Government Social Research**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Aim of the research is to gather information from professionals specialising in child sexual exploitation. **Is the study clear in what it seeks to do?** Clear. Aim to scope nature of CSE in Scotland. **How defensible/rigorous is the research design/methodology?** Defensible. A practice seminar including focus groups and opportunity for individual feedback took	**Is the context clearly described?** Clear: 27 practitioners from a range of agencies across Scotland, representing child protection committees, health, the police, and third sector organisations. **Was the sampling carried out in an appropriate way?** Not Sure. Not reported. **Were the methods reliable?** Somewhat reliable. One method -	**Does the study's research question match the review question?** Study explores key professionals within the field of CSE, and relevant findings are reported to include inter-agency arrangements and the effectiveness and hindrances. **Has the study dealt appropriately with any ethical concerns?** No. Not reported.	**Overall assessment of internal validity:** - Due to no reporting on the data collection or analysis of the seminar, findings are not rich. There is no contextualising of participants or ascribing which finding was said by who. Consequently, conclusions are somewhat adequate. **Overall assessment of external validity:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
place in Edinburgh inviting practitioners from a range of agencies across Scotland - the purpose of the day was to understand the scale and nature of CSE and professional perception. **How well was the data collection carried out?** Not sure. Not reported.	One day event for focus group, individual questionnaire and general discussion. **Are the data 'rich'?** Not sure. Not adequately reported how findings were drawn, or which professional reported, therefore difficult to contextualise. **Is the analysis reliable?** Not sure. Not reliable. **Are the findings convincing?** Somewhat convincing. The findings are supported by a thorough literature review. At the beginning of Chapter 6, the authors report that 'the views expressed are not necessarily representative, and to this extent should be treated with some caution'. The purpose is to inform the nature of CSE throughout Scotland, and highlight further research needs. **Are the conclusions adequate?** Somewhat adequate. Key messages are drawn from 4 key themes highlighted in chapter 6 which report the practitioners who attended the seminar reporting a varying level of awareness, knowledge and experience of CSE	**Were service users involved in the study?** No. Report not co-produced. **Is there a clear focus on the guideline topic?** Yes. CSE. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. CSE. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Child protection professionals. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. Child protection professionals working within CSE to explore the 'nature and extent of inter-agency communication and working in relation to child sexual exploitation; and the barriers to the development of effective practice' (p13). **Does the study have a UK perspective?** Yes. Scotland.	Aim is relevant to research question. No consideration of ethical approval. **Overall validity rating:** -

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	across Scotland. Findings highlight these gaps in knowledge and there is a recognition by authors for an improvement in partnership working and the creation of better systems for sharing information.		

5. **Crockett R, Gilchrist G, Davies J et al. (2013) Assessing the Early Impact of Multi Agency Safeguarding Hubs (MASH) in London. London: London Councils, MASH and University of Greenwich**

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Thorough justification of interviewing MASH staff, with baseline and follow-up data collection. **Is the study clear in what it seeks to do?** Clear. Aim is defined for phase 3 as: 'pre implementation interviews was to gather a broad range of opinion on the aims, and expected outcomes of the programme, and also on how the programme is being implemented in its early stages. The post implementation interviews were conducted to explore the impact of the introduction of the MASH on the work of individual professionals and on safeguarding more generally. The post implementation interviews also aimed to capture some of the changes that	**Is the context clearly described?** Clear. Clear characteristics of participants who represent each agency in both pre and post implementation interviews. Additional information is given for why the pre (n=24) interviews and the post (n=16) is. Additionally, interwoven throughout the findings, it is generally ascribed to which professional said what. **Was the sampling carried out in an appropriate way?** Appropriate. Through manager, and then research team contacted respondents to represent an array of MASH professionals. Limitations are not discussed as to potential for bias.	**Does the study's research question match the review question?** Yes. Analysis of Multi-Agency Safeguarding Hubs (MASH) with baseline and follow up qualitative interviews with staff in order to consider the early impact. **Has the study dealt appropriately with any ethical concerns?** Yes. Participant consent forms used and attached in Appendix 2. **Were service users involved in the study?** No. Study has not been co-produced. **Is there a clear focus on the guideline topic?** Yes. Child protection.	**Overall assessment of internal validity:** ++ Comprehensive empirical study with clear aims and findings that are collected at two stages. **Overall assessment of external validity:** ++ Analysis of MASH and the early impact on staff in pre and post implementation qualitative study. **Overall validity rating:** ++ Good study however caution to generalise as findings are representative of 5 boroughs in London.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
the move to MASH had brought to safeguarding services' (p37). **How defensible/rigorous is the research design/methodology?** Defensible. Comprehensive methodological section which details why boroughs were selected and the use of purposive sampling to conduct interviews at two data collection points. **How well was the data collection carried out?** Appropriately. Telephone interviews were conducted with a range of MASH professionals both prior to the implementation of the MASH and approximately 2 months later. Participants were identified through their manager and contact details sent to the research team. Interviews are semi-structured using an interview schedule developed in line with the study aims (Appendix 3 attached).	**Were the methods reliable?** Reliable. Study includes 4 phases of data collection. Phase 3 is only relevant to Q21. **Are the data 'rich'?** Rich. Context is clearly described and a variety of professionals contribute to the findings, as detailed by the research team who ascribes individual comments to interviewees and how many mentions (through anonymous references i.e. MP). There is a balanced presentation of findings. **Is the analysis reliable?** Reliable. Framework for analysis developed by Richie and Spencer (1994) - with how the research team analysed findings. **Are the findings convincing?** Convincing. Clearly presented into inductive themes. Data is analysed thoroughly and follows a framework matrix (Richie & Spencer 1994). **Are the conclusions adequate?** Adequate. Summary by authors presents an overview of challenges and enablers as described by interviewees.	**Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Child protection agencies - police, health, CSC staff, education, probation and housing. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. Multi-agency safeguarding hubs. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Child protection. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. Early impact of MASH by collecting pre and post implementation qualitative interviews with MASH professionals. **Does the study have a UK perspective?** Yes: 5 London boroughs.	

6. Mortimer J, North M, Katz A at al. (2012) You have someone to trust - Outstanding safeguarding practice in primary schools. London: Office of Children's Commissioner

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Mixed. Main focus on safeguarding practice with some data on multiagency working. **How defensible/rigorous is the research design/methodology?** Defensible. Online survey (Qualitative), focus groups, semi-structured interviews of professionals in educational settings (primary schools). **How well was the data collection carried out?** Somewhat appropriately. Via interviews, focus groups and online survey, no details reported on how participants were recruited.	**Is the context clearly described?** Clear. Primary school, educational context. **Was the sampling carried out in an appropriate way?** Not sure. No details reported on how participants were recruited. **Were the methods reliable?** Reliable. **Are the data 'rich'?** Mixed. For views of the educational professionals. **Is the analysis reliable?** Not sure/not reported. No details reported on how data were analysed. **Are the findings convincing?** Somewhat convincing. Unclear methods of data analysis. **Are the conclusions adequate?** Not sure. Due to limited details reported.	**Does the study's research question match the review question?** Yes. To identify best professional practice in response to child protection and safeguarding concerns in primary schools, including both in-school practice and interagency working from the school's perspective. **Has the study dealt appropriately with any ethical concerns?** Partly. Consultation with selected school staff, external partners, with education professionals and wider group of schools regarding interviews and focus groups. No details on ethical approval/consent. **Were service users involved in the study?** Yes. Participants involved as subjects in this study. **Is there a clear focus on the guideline topic?** Partly. Main focus is on safeguarding in schools,	**Overall assessment of internal validity:** - Lack of methodological details. **Overall assessment of external validity:** + **Overall validity rating:** -

		with some data on best professional practice relating to multiagency working. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. School staff, education professionals and wider school staff and external partners. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. School environment. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Safeguarding with some data on multiagency working. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. Relating to interagency working, the challenges in maintaining these relationships with other agencies. **Does the study have a UK perspective?** Yes.	

7. Rouf K, Larkin M, Lowe, G (2012) Making decisions about parental mental health: an exploratory study of Community Mental Health Team staff. Child Abuse Review 21: 173–89

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Defensible. **How well was the data collection carried out?** Appropriately. Semi-structured interviews and diaries.	**Is the context clearly described?** Clear. **Was the sampling carried out in an appropriate way?** Somewhat appropriate: 99 Community Mental Health Team (CMHT) staff were contacted and 13 staff participated in the study (low response rate and participants likely to be self-selective). **Were the methods reliable?** Reliable. Diaries observations were triangulated with interview data. **Are the data 'rich'?** Mixed. A mixture of data, some not directly related to interagency working. **Is the analysis reliable?** Reliable. All data, were subjected to reliability checks, transcribed verbatim and systematically analysed using Interpretative Phenomenological Analysis (IPA), with detailed reading and re-reading of material on a case-by-case basis for themes to	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Yes. Ethical approval granted by the local Research Ethics Committee and consent from participants obtained. **Were service users involved in the study?** Yes. Professionals as participants in study. **Is there a clear focus on the guideline topic?** Yes. To identify and explore themes relating to tensions of working within and across systems, trying to balance perceptions and feelings involved in decision-making, and the role of interpersonal dynamics in the understanding and management of risk. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Mental health professionals.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** ++ **Overall validity rating:** + The methodology was sound but the poor response rate was low.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	gradually emerged within interviews, and overarching themes identified. **Are the findings convincing?** Somewhat convincing. Only part of the findings relates to data on interagency working. **Are the conclusions adequate?** Somewhat adequate. Only part of the findings relate to data on interagency working.	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Parental mental health and child welfare. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. Part of the views expressed relate to the tension of interagency working. **Does the study have a UK perspective?** Yes.	

8. Smeaton E (2013) Running from hate to what you think is love: the relationship between running away and sexual exploitation. Ilford: Barnardo's

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Qualitative comp 1 Which component? Consultation with young people with experience of running away and CSE while under the age of 16.	**Quantitative component descript A (including incidence or prevalence study without comparison group; case series or case report)** Which component? Consultation with professionals working with children and young	**Does the study's research question match the review question?** Yes. Study explores professional practice in relation to both recognition (Q6) and response (Q20). Additional findings can relate to Q.21 as the study explores multi-agency working.	**Overall assessment of external validity** - **Overall assessment of internal validity** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question? Yes. **Is the process for analysing qualitative data relevant to address the research question?** Yes. Thematic analysis of each young person's story, then second-stage thematic analysis of common themes. **Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. Little consideration given to variation in the experiences of young people by gender, age, ethnicity and so on. Also unclear how their localities impacted their experience. **Is appropriate consideration given to how findings relate to researchers' influence; for example, though their interactions with participants?** No.	people experiencing CSE (survey). **Is the sampling strategy relevant to address the quantitative research question (quantitative aspect of the mixed-methods question)?** Partly. Respondents recruited via the NWG Network's newsletter. No consideration of possible gaps that this strategy might leave. All organisations were voluntary sector organisations (although 2 were based within the statutory sector) - it is not clear in the methodology whether this was intentional or simply a product of who responded. **Is the sample representative of the population under study?** Partly. Poor response rate - 28 organisations responded from 500 contacts - suggesting possible bias in favour of most motivated respondents. **Are measurements appropriate (clear origin, or validity known, or standard instrument)?** N/A. Qualitative questions rather than measures.	**Has the study dealt appropriately with any ethical concerns?** Partly. Study was scrutinised and approved by the Barnardo's Research Ethics Committee. One of the principles of the research is given as 'ensuring informed consent and assent' (p12), however no details are given about how consent was obtained. Consideration is given to any needs/issues raised during the research and how these will be addressed. **Were service users involved in the study?** No. Service users involved as participants, but not in conducting or analysing the research. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Children and young people at risk of or who have experienced sexual exploitation, and professionals working with sexually exploited young people.	Could be more detailed description of how addressed ethical issues. **Overall score** - Survey of services is entirely of voluntary sector services, and it is unclear whether interviewed professionals represented a wider range of services - the voluntary sector perspective of the research is not highlighted or justified in the research methodology. Little consideration in the findings of how contextual and demographic factors shape participant responses.

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Qualitative comp 2 Which component? Consultation with professionals working with children experiencing CSE (telephone interviews). **Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question?** Partly. Professionals working with young people experiencing running away and CSE appropriate. However, unclear if all relevant sectors were represented - job titles suggest that they were mainly in the voluntary sector. **Is the process for analysing qualitative data relevant to address the research question?** Yes. Two-stage analysis. **Is appropriate consideration given to how findings relate to the context, such as the setting, in which the data were collected?** No. Little consideration of how different localities, types of service and so on affected findings. **Is appropriate consideration given to how findings relate to**	**Is there an acceptable response rate (60% or above)?** No. Response rate appears to be 5.6% (28 responses from 500 contacts) which is poor.	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to Q6 and Q20. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **g. Does the study have a UK perspective?** Yes. England.	

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
researchers' influence; for example, though their interactions with participants? No.			

9. Taylor J, Stalker K, Fry D et al. (2013) Disabled children and child protection in Scotland: investigation into the relationship between professional practice, child protection and disability. London: Scotland. Scottish Government Social Research

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. **Is the study clear in what it seeks to do?** Mixed. Study's research questions are clear, but findings are not written up according to the original research questions. **How defensible/rigorous is the research design/methodology?** Somewhat defensible. Participants had to have experience of at least two cases of child protection concerns in relation to a disabled child, which would seem to be relatively low. The spread of professional backgrounds and extent of experience of recruited participants is not reported. **How well was the data collection carried out?** Somewhat appropriately. Use of Critical Incident	**Is the context clearly described?** Unclear. Little consideration of how respondents' roles and experience may affect their responses. **Was the sampling carried out in an appropriate way?** Not sure. Insufficient information to judge. **Were the methods reliable?** Not sure. Insufficient information regarding focus groups. **Are the data 'rich'?** Mixed. Use of case studies adds to richness of data, but overall there is little consideration of reasons for divergences in opinion between interviewees. **Is the analysis reliable?** Somewhat reliable. Results analysed in various ways, including use of modelling. With regard to thematic analysis of qualitative data it is not	**Does the study's research question match the review question?** Yes. **Has the study dealt appropriately with any ethical concerns?** Partly. Ethical approval by University of Edinburgh Moray House School of Education Ethics Committee. Mentions informed consent for interview participants, but not focus group participants. **Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Practitioners working with disabled children at risk of or experiencing significant harm.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Limited information on ethical considerations in relation to focus groups. **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Technique as part of in depth interviews. All interviews digitally recorded with consent. Less clear how focus groups were conducted.	clear how themes were formulated and checked, and some of the themes identified do not appear to be coherent concepts. **Are the findings convincing?** Somewhat convincing. Themes developed in thematic analysis do not appear to be coherent and distinct. **Are the conclusions adequate?** Somewhat adequate.	**Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study relates to recognition and response. **(For views questions) Are the views and experiences reported relevant to the guideline?** Yes. **Does the study have a UK perspective?** Yes. Scotland.	

10. Vincent S and Petch A (2012) Audit and Analysis of Significant Case Reviews. Edinburgh: Scottish Government

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
Is a qualitative approach appropriate? Appropriate. Content analysis of SCR reports. **Is the study clear in what it seeks to do?** Clear. **How defensible/rigorous is the research design/methodology?** Defensible. **How well was the data collection carried out?** Appropriately. Relevant content extracted from reports using template.	**Is the context clearly described?** Clear. Limitations of SCRs as a source of data made clear. **Was the sampling carried out in an appropriate way?** Appropriate. All SCRs published in a particular timeframe (post 2007). **Were the methods reliable?** Somewhat reliable. Template used for analysing reports. However, no mention of double coding	**Does the study's research question match the review question?** Yes. Partly, overall research question is about learning from Serious Case Reviews, but there is one section relating to assessment. **Has the study dealt appropriately with any ethical concerns?** Partly. Steps taken to ensure that information from SCR reports remained anonymised. No mention of ethical approval.	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Information on assessment is part of a broader study. **Overall validity rating:** +

Internal validity – approach and sample	Internal validity – performance and analysis	External validity	Overall validity rating
	or cross-validation by a second member of the team. **Are the data 'rich'?** Mixed. Little contextualisation of findings in the context of cases. Some direct quotes from SCR reports used. **Is the analysis reliable?** Not sure/not reported. Little data presented on which to base this judgement. **Are the findings convincing?** Somewhat convincing. Little presentation of primary data to show how particular themes/issues have been identified. **Are the conclusions adequate?** Somewhat adequate.	**Were service users involved in the study?** No. **Is there a clear focus on the guideline topic?** Yes. **Is the study population the same as at least 1 of the groups covered by the guideline?** Yes. Study relates to cases where children have died or been injured, the majority of which had an element of abuse or neglect. **Is the study setting the same as at least 1 of the settings covered by the guideline?** Yes. **Does the study relate to at least 1 of the activities covered by the guideline?** Yes. Study as information relevant to assessment. **Does the study have a UK perspective?** Yes. Scotland.	

Review question 21 – Findings tables

1. **Beckett H, Brodie I, Factor F et al. (2013) 'It's wrong … but you get used to it' - A qualitative study of gang-associated sexual violence towards, and exploitation of, young people in England. London: Office of the Children's Commissioner for England**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: 'The research aimed to consider: the scale and nature of gang-associated sexual violence and exploitation in six areas of England; the main pathways into gang-related sexual violence and exploitation for young people living in these neighbourhoods; and potential models for an effective multi-agency response to the issue' (p6). **Methodology:** Qualitative study. The research team adopted a qualitative approach to conduct: - Individual interviews with 150 young people - 11 focus groups with 76 professionals - 8 single sex focus groups with 38 young people There is a comprehensive methodology section (pp12–15) that addresses approach to	**Participants** Children and young people. Individual interviews - Young people aged 13-28 (n=150). Focus groups - Young people (n=38). Professionals/practitioners - 11 focus groups were conducted with 76 professionals across 6 research sites. Representation from fields of social care, education, health, policing and the justice system, specifically working within the gangs and sexual exploitation/sexual violence. **Sample characteristics** • Age - Interviews - Participants ranged from 13 to 28: Under the age of 18 (49%); 18-20 (28%); 21–25 (21%); and 25–28 (2%). Focus groups - Not reported. Professionals - Not reported. • Sex - Interviews - 52% were male, with 48% female. Focus groups - Not reported. Professionals - Not reported. • Ethnicity - Interviews - The self-reported ethnicity of interviewees: 32% Black/Black British; 28% White; 21% Dual heritage; and 18% Asian/Asian British. Focus groups - Not reported. Professionals.	**Narrative findings** Q21 relevant data Professional's and young people's view on multiagency working showed that gang-associated sexual violence and exploitation still in the early stages of development in terms of how it prevents, identifies and responds. The overall assessment of professional engagement was still primarily reactive, as opposed to proactive and preventative. Common factors repeatedly identified as conspiring against both the effective identification of, and an effective response to, the issue of gang-associated sexual violence and exploitation: a. A lack of trust in the police and the criminal justice system by young people 'We don't believe that police are there to help us … I've been hit on my leg with a kosh, it's not necessary' (young women) (p45). b. A lack of information flow to and from many gang-affected neighbourhoods and a historical lack of integration between community-based and statutory based initiatives attempting to tackle issues within the community. c. Insufficient awareness, acknowledgement of and engagement with the risks of gang-associated sexual violence and exploitation amongst many different professionals - too much focus on young men and male-on-male physical violence within gangs work and a neglect of female and/or sexually based gang victimisation (p46).	**Overall assessment of external validity** ++ The study meets all criteria and has dealt effectively with ethical considerations. In addition, the interview guide was co-produced with the YPAG to make questions age appropriate. **Overall assessment of internal validity** ++ Very comprehensive, effective study with detailed findings that are summarised into recommendations. The methodology has dealt appropriately with the ethics and risk associated with the subject matter, and the research team have ensured the voice of the child is at the heart of the report. **Overall score**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
interviewing; the research was co-produced with young people from the Young People's Advisory Group; information on gaining consent; and thorough steps taken to provide confidentiality and anonymity of participants. The study took place between 2011 and 2013. **Country:** UK, England. 'To maintain confidentiality and protect participants, the identity of the research sites is not being revealed' (p6). **Source of funding:** Government - Inquiry of the Office of the Children's Commissioner into child sexual exploitation in gangs and groups. Led by the University of Bedfordshire.	• Religion/belief - Interview - Not reported. Focus groups - Not reported. Professionals - Not reported. • Disability - Interview - Not reported. Focus groups - Not reported. Professionals - Not reported. • Long term health condition - Interview - Not reported. Focus groups - Not reported. Professionals - Not reported. • Sexual orientation - Interview - Not reported. Focus groups - Not reported. Professionals - Not reported. • Socioeconomic position - Interviews - 'Most participants reported that they were in some form of education (45%), training (20%) or employment (18%), with only one in eight identifying as Not in Education, Employment or Training (NEET). Focus groups - Not reported. Professionals - Not reported. • Type of abuse - 87% (n=131) had direct, often multiple connections with gangs. of the 131 participants, 59% were/had been directly involved in a gang (M=70% v. F=47%); 32% had been gang-associated (M=25% v. F=39%); 35% had friends/and or family involved; 23% were having/had previously had a 'romantic relationship; with a gang-involved person (all female bar one); 57% had personal experiences of sex and/or relationships in gangs. The remaining 13%	d. Silo working across different agencies and strands of work and a lack of knowledge of how one's practice fits within wider relevant strategies and operational initiatives. e. A consequent lack of information-sharing between agencies. f. Inadequate partnership-working and cross-fertilisation of learning between gangs and sexual exploitation/sexual violence initiatives (p46). g. Lack of clarity as to the best policy fit for these issues: is it best conceptualised and responded to as child sexual exploitation, domestic violence, violence against women and girls (VAWG) and/or serious youth violence? h. Impact of financial cuts on the provision of services, both in terms of which services remained and decreasing capacity to engage in any long-term supportive work, a key to any sustainable response. Under-resourcing hampered a planned inter-agency systemic response long term (p48). '... young people don't understand the fact that things are commissioned or funded for a set period of time … – this week we're running; next week we're not running ever again because our funding's finished' (professional) (p48). 'Without funding it's not gonna change. It's not gonna change. It's a waste of time. They can have these debates but there's no money to fund changes' (young people) (p48). j. Inadequate data collection and data monitoring - risks may be being picked up, but not being systemically recorded or monitored. There is little read across between these databases (p47).	++ A thorough empirical study which meets its research aim and details implications for practice and policy on a local and national level.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	(n=19) participants grew up in gang-affected neighbourhoods. Focus groups - Not reported. Professionals - Not reported. • Looked after or adopted status - 38% of participants reported current or previous involvement with children services, although it is not clear what support this was. Focus groups - Not reported. Professionals - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Interviews - Not reported. Focus groups - Not reported. Professionals - Not reported. **Sample size** Interviews - 150 participants Focus groups - 8 single sex with 38 young people Professionals - total of 11 focus groups held with 76 professionals.	A number of principles central to effective preventative initiatives was identified by participants (both professionals and young people) (p50). These included: a. Sustained co-investment in universal and targeted preventative work. b. Active school engagement in preventative efforts. c. Commencing preventative work at an early age (primary school level). d. Using 'credible' individuals to deliver preventative messages. e. Supporting parents/carers to identify and respond to risk. f. Engaging the wider community in preventative initiatives. g. Engaging young people as partners in identifying solutions, at both an individual and systemic level (p50).	

2. **Berelowitz S, Clifton J, Firmin C et al. (2013) 'If only someone had listened': Office of the Children's Commissioner's inquiry into child sexual exploitation in gangs and groups. London: Office of the Children's Commissioner for England**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: Based on Phase 2 of the inquiry - To assess views and experiences of victims of and professionals	**Participants** Professionals/practitioners – 74 Children and young people – 15 Parents/Caregivers - 11	**Narrative findings** Q21 relevant data This study, based on Phase 2 of an inquiry, assessed views of children and young people of CSE and professionals working in CSE found 9 system failings	**Overall assessment of internal validity:** + **Overall assessment of external validity:**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
working in CSE: 2 objectives: • To learn where and how child sexual exploitation is already being tackled successfully, i.e. Why CSE children continue to be let down • To understand what's getting in the way – the 'barriers' − where the problem is not being dealt with effectively. To assess views and experiences of victims of and professionals working in CSE, i.e., Why there is a no fully joined-up multi-agency, child-centred approach to address child sexual exploitation in gangs and groups, why agencies and individuals fail to listen to them, and fulfil their responsibilities with regard to child protection, or that there was not sufficient strategic and managerial oversight to coordinate their actions.	**Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** Based on Phase 2 of this Inquiry Workshops, interviews and seminars (total 23) with children and young people, professionals and academics: Gather evidence from presentations, group discussions and interviews: children/young people (n=15); parents/carers (n=11); professionals (n=74); academics (n=11) (Fig 1, p18).	and challenges in how agencies work together to tackle CSE: 1. Child's best interests the top priority Many agencies forgetting the child, many children and young people are being lost or overlooked by the system. 'They talked about me like I wasn't even there. They were very harsh about me' (p22). Example of good practice: professionals be given the time and space to focus on supporting children and young people, and focus on their individual needs and equalities; conflicting priorities impeded effective practice. 2. Gaining the child's confidence and participation of children and young people in decision-making Services failing to engage with children and young people. Professionals failed to understand, recognise and accommodate to their individual needs, language, beliefs and feelings. Support be tailored to meet the needs of the child, 'They didn't even ask me if I was OK or if it's OK to talk about it' (p24). Example of good practice: building an informed and supportive environment that enabled children and young people to have the confidence to come forward to talk to professionals when they are worried. 3. Leadership There was a lack of clear and committed leadership amongst some of the most senior decision makers at local level. Without local and national leadership, dedicated professionals worked in a vacuum. Example of good practice: adopting and ensuring a 'whole-school	++ **Overall validity rating:** + Overall methodology of this study was sound though some details missing (sampling and population characteristics). Data analysis was robust and interpretations of findings convincing.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Methodology: Based on Phase 2 of the Inquiry semi-structured interviews and focus groups with children, young people, parents and carers; site visits and workshops and academic seminars. **Country:** UK. **Source of funding:** Government. Office of the Children's Commissioner.		approach' to protecting children and young people in schools (p24). 4. No strategic planning in some LSCBs in relation to CSE. There was an absence of a joint strategy results in differing approaches and conflicting priorities between local agencies. Example of good practice: to formulate a national action plan to tackle CSE and oversee its delivery to bring greater consistency to service delivery and local planning (p26). 5. Everyone on alert Too many people who should be protecting children were in denial about the realities of CSE despite the mounting public, political and media interest in child sexual exploitation. 'People should have thought of these questions …' (p26). Supervision, support and training of staff needed investing. Example of good practice: to raise the awareness of communities, professionals and children and young people through campaigns and training. Investment in supervision, support and training of staff (p41). 6. Spotting the warning signs Professionals failing to recognise victims. There was patchy understanding of child sexual exploitation around the country, and prejudices that prevented professionals from recognising both victims and perpetrators when they did not conform to their preconceptions (p28). Example of good practice: awareness-raising and training for multi-agency professionals.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		7. Coherent joined-up working and effective information-sharing within and between agencies Various agencies and services working in isolation to tackle CSE and viewed child sexual exploitation through its own lens. They failed to work together to arrive at a comprehensive picture of the problem in their local area. These difficulties were compounded by limited and incompatible IT systems which impede information-sharing and effective communication between agencies (p29). Example of good practice: The establishment of a multi-agency forum (such as the Multi-Agency Safeguarding Hubs (MASH)) to combine the expertise and resources of several bodies in order to identify and refer children and young people who are at risk of child sexual exploitation, as some sexually exploited children and young people face dangers from multiple sources (p46). 8. Pre-emptive action Taking pre-emptive action to break up networks that exploit children and to avoid delayed response to CSE, far more effective approach than waiting until a child reveals that he or she is being abused. These delays were further aggravated in some areas by bureaucratic and time consuming processes often driven by agency procedural requirements (p30). Example of good practice: Police forces worked in partnership with housing, sexual health, social care, domestic abuse and missing children's services, alongside anti-social behaviour teams and schools, to compile intelligence on CSE and then take steps to stop it happening (p51).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		9. Scrutiny and oversight, evaluation and review Results not being monitored. Statutory agencies failing to check whether their actions were working and there was no common agreement between them as to what they were trying to achieve (pp30–1). Example of good practice: sufficient resources given in terms of funding.	

3. **Brandon M, Sidebotham P, Bailey S et al. (2013) New learning from serious case reviews: a 2-year report for 2009 to 2011. London: Department of Education**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To identify common themes and trends across the 2009–11 review reports on child maltreatment and neglect, drawing out the implications for policy and practice; to provide up to date comprehensive data on fatal maltreatment of children in England and to set these in the context of other relevant data on children's health, well being and possible harm. **Methodology:** Qualitative Study. Serious	**Participants:** **Sample characteristics:** • Age - Mostly on children aged 5–10 years. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported.	**Narrative findings** Common themes in relation to services, professional attitudes, knowledge and behaviours, and the systems and structures that underpin safeguarding: 1. A culture of procedure-driven, uncritical practice in teams can contribute to 'silo practice' and side-lining/exclusion of different professionals in a few cases, which may have arisen because of professionals focusing exclusively on their own areas of practice, again taking a narrow, problem-based approach to working with children and families (p80); or due to different understandings of criteria and thresholds for provision (p78). In some cases there was evidence of fragmentation of adult services, for example between alcohol services and other mental health services (p82). 2. A lack of professionalism and critical thinking among practitioners, not taking their safeguarding roles seriously, leading to individuals passing the	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + **Overall validity rating:** + This serious case review analysed national data based on a large sample of cases, with fairly sound methodology.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
case reviews of 5 individual and interlinking studies into child maltreatment and neglect. **Country:** UK. **Source of funding** Voluntary/charity - Grant from the National Institute for Mental Health, with support from the Violence Prevention Branch of the US Centres for Disease Control and Prevention.	**Sample size:** The overall analysis includes 5 interlinking studies drawing primarily on either the 115 serious case reviews notified to the Department for Education during the single year 2009–10, or the full sample of 184 serious case reviews from the two year period 1 April 2009–31 March 2011.	buck, or relinquishing their responsibility once they had referred the case on to others and not ensuring that actions did take place. This was reflected in the issues around incident-driven practice, the rule of optimism and failure to consider the child's perspective. This lack of professionalism could extend to the underlying culture of whole teams, resulting in inadequate assessments, or a failure to follow cases through from assessment to actions and outcomes. A lack of professional approach and critical challenge within teams can also extend to supervision (pp81–2). 3. An over-reliance on electronic recording systems and proformas, and working strictly to criteria rather than critically thinking about cases (p83). 4. Professionalism and critical approaches to practice require both training and experience, and systems that support such approaches. 5. Inexperienced and newly qualified practitioners dealing with complex cases, whereas professionals with the most experience tend to be in managerial/supervision roles and have very little direct contact with children and families. Systems of peer supervision needed to be developed (p83). 6. Inter-agency working takes time in liaising with others, following through on actions, and challenging and escalating when necessary. Critical reflection, peer review and supervision all require adequate time if they are to be effective. Too often professionals are driven by the needs of the system, and do not take the time to stop and think.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		7. Safeguarding children is demanding work that takes its toll on practitioners, there needs to be structures for appropriate support of front-line workers to lessen the impact of this work (p84). 8. The child protection systems a. Confusion among practitioners between 'child in need' procedures and 'child protection' procedures as a continuum, leading to a substantial gulf in practitioners' approaches. There was also confusion over the terminology used for multi-agency meetings, including 'child in need', 'common assessment framework', and 'team around the child' meetings, compounded by a lack of clarity in terms of who takes responsibility for such meetings, lack of clear arrangements for chairing and taking minutes, and a lack of structure for the meetings. This led to many meetings being unclear in their focus, with a lack of any definitive action plan or accountability for following through on agreements, resulting in inadequate assessments being undertaken or repeated partial assessments which never fully appraised the situation of the children (pp84–5). b. Inter-agency working and involvement of the courts - there were significant difficulties and barriers to involvement between court processes and inter-agency working to safeguard children. Court proceedings were seen as separate from inter-agency working. It appeared that court decisions were affecting ability of professionals to continue safeguarding work. Misunderstandings of these processes and breakdowns in communication may lead to children being put at further risk of harm. There is need for further research and consultation into how the courts and other agencies work together to effectively safeguard and promote the welfare of children (p88).	

4. **Brodie I, Pearce J (2012) Exploring the scale and nature of child sexual exploitation in Scotland. Edinburgh: Scottish Government Social Research**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The study had 3 aims: 'to review existing research, policy and practice literature from the UK regarding the scale and nature of CSE, and trafficking for sexual exploitation, focusing on Scotland ... (gather) exploratory information from key professionals regarding their perceptions of the scale and nature of CSE in Scotland' (p3). **Methodology:** Qualitative study. Focus group held with 27 practitioners. **Country:** UK, Scotland. **Source of funding:** Government - Scottish Government.	**Participants** Professionals/practitioners - 27 practitioners' representative of Child Protection Committees, health, the police and third sector organisations. **Sample characteristics** • Age – Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** n=27.	**Narrative findings** (Findings relevant from report are Chapter 6 - The expert seminar) Questionnaire findings: - The police were identified by almost all practitioners (n=21) as the most active agency, followed by third sector and children's services respectively (p41). - Issues identified by practitioners about information were that practice around this issue was varied. -Two practitioners responded highlighting that police operations have stimulated efforts to develop a more multi-agency, information sharing efforts (p41). Focus group findings are reported in 4 overarching themes which were prevention, identification, support, and disruption and prosecution. For the purpose of the research question, organisational factors, relevant themes are support, and disruption and prosecution. 6.21–6.25 Support - By and large, practitioners understood specialist services are run primarily by the third sector, it was felt that provision for sexually exploited young people was inconsistent. Practitioners did not elaborate on what additional provision would be like, potentially as expertise had not developed in this area (p43). - There was some contradiction about where provision ought to be located for sexually exploited children and young people. However some commented	**Overall assessment of internal validity:** - Due to no reporting on the data collection or analysis of the seminar, findings are not rich. There is no contextualising of participants or ascribing which finding was said by who. Consequently, conclusions are somewhat adequate. **Overall assessment of external validity:** + Aim is relevant to research question. No consideration of ethical approval. **Overall validity rating:** -

		on the significance of statutory services being fit to respond to sexual exploitation, for example Looked after Children (p43). - Raising awareness and gaps in training were identified as an issue because of professionals concerns over their expertise to respond to sexual exploitation. - Most practitioners concerns were communicated about the number of provisions for over 16s, including care leavers. The connections amongst the link between adult and children's services were likewise felt to be tricky – reflected, for instance, in various thresholds enforced by provisions to supporting children, young people and adults (p43). 6.26–6.31 Disruption and prosecution A small number of practitioners had experience gathering evidence, sharing intelligence and working with police to disrupt abusers (p44). Issues cited: - The existing legislation did not aid in convicting abusers. - Perceived gaps in the system, for example, if 'practitioners were not sufficiently trained or supported, then the disruption of abusive networks would not take place' (p44). - areas for development in practitioners knowledge were identified – e.g., 'how the process of gathering evidence should take place, and how witnesses could best be supported. Practitioners were keen to find out more about what had been learned from recent police operations in Scotland' (p44). Potential areas for development were recognised by practitioners to be where agencies work alongside the police, to assist in gathering information and recognising local patterns of abuse.	

		Practitioners felt that work could be developed in all areas – prevention, identification, support and disruption. Improved partnership working and better systems for sharing information were viewed as important elements of developing good practice (p45).	

5. **Crockett R, Gilchrist G, Davies J et al. (2013) Assessing the Early Impact of Multi Agency Safeguarding Hubs (MASH) in London. London: London Councils, MASH and University of Greenwich**

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To 'examine the effect MASH has had on supporting practitioners in delivering effective and focused interventions, and furthermore changing approaches to safeguard practice' (p15). **Methodology:** Qualitative study. Qualitative interviews with MASH staff only relevant (phase 3) but data collection has other phases: interviews with key stakeholders; visits to MASH; observational data to consider physical set up of rooms; and administrative data on referrals. **Country:** UK.	**Participants:** Professionals/practitioners. Phase 3 of interviews: - Phase 3 (pre-implementation) included qualitative interviews with 24 multi-agency practitioners from each MASH to represent the range of disciplines involved in each MASH: social work (n=7); health (n=5); police (n=5); education (n=2); probation and yot (n=5); and housing (n=1). Phase 3 (post-implementation) 16 interviews with multi-agency practitioners: social work (n=4); health (n=4); police (n=3); education (n=2); probation and YOT (n=3); and housing (n=0). **Sample characteristics:** • Age – Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported.	**Narrative findings** For the purpose of research question 21, Phase 1 (Observation of MASH) and Phase 2 (Audit of administrative records of referrals) and Phase 3 (Interviews with MASH referrers) are not applicable. Chapter 6 explores Phase 3 findings where numbers relate to the relevant section and themes as indicated in the paper. Where there are gaps in numbering it should be noted that the section or theme is not relevant to Q21. Pre implementation interview findings are: 1. Communication and information sharing - Interviewees recognised the importance of sharing information, e.g. 'every single serious case review talks about information not being shared' (MP8, p 40). - There were numerous issues cited by interviewees: challenges in knowing who the right person is (e.g. which health visitor to contact); staff absence leaving gaps and delays in response; and in certain cases, parental consent being a necessity before information is shared.	**Overall assessment of internal validity:** ++ Comprehensive empirical study with clear aims and findings that are collected at 2 stages. **Overall assessment of external validity:** ++ Analysis of MASH and the early impact on staff in pre and post implementation qualitative study. **Overall validity rating:** ++ Good study however caution to generalise as findings are representative of 5 boroughs in London.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Source of funding Government. London Councils.	• Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not relevant. **Sample size:** Phase 3 (t=40) qualitative interviews with MASH professionals. **Follow up:** Phase 3: Pre-implementation and 2 months post-implementation.	- Professionals cited the Seven Golden Rules to information sharing, the Pan London Information Sharing Agreement and protocols specific to each agency. - Parental consent was a worry for some professionals who highlighted the recent case in Haringey, where parents successfully sued a local authority for sharing information without consent. - GPs (who are external to MASH) were reported frequently for being reluctant to share data and respond to requests for information. - Not having access to databases and all the various databases used by different professionals was a frustration for some interviewees. 2. Roles and inter-professional working 2.1 Different cultures - Some interviewees commented upon the contrasting way other agencies respond to safeguarding concerns. For example, 1 police officer (MP24, p41) described himself and colleagues 'as being trained to make rapid decisions and compared this to social workers who take a more "softly softly" approach that takes longer'. Other interviews described the police as 'having their own way of doing things' (MP10, p41). - Seven interviewees felt that MASH was or would facilitate positive working relationships, despite the different professional cultures, and would foster a better understanding of roles and responsibilities. - For Lewisham borough (where MASH had been live for the longest) 1 respondent commented that prior to MASH, agencies would only meet to discuss cases if there was a disagreement but now working together, decisions were made more accurately and timely.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		2.2 Working together - Co-location was seen to be promoting relationship building, mutual professional understanding and the development as trust. One interviewee commented 'having professionals in one room, you establish a level of trust, understanding which may not have been quite as strong when you're all in separate areas' (MP5, p42). - Having all agencies in one secure space was seen to be saving on traveling times. - A couple of interviewees felt cut off from their individual agencies if they were co-located with MASH teams. - Challenges cited: agencies having different risk thresholds; language and terminology; working styles and cultures. In one instance, a police officer commented on the 'hierarchical, disciplined nature of the police force had become accustomed to working in an open plan office with a different management style at MASH, but has found this an interesting experience rather than a difficulty' (p42). Post implementation interview findings are: 1. Information 1.1 Communication - MASH was seen to facilitate better communication, which ensured high quality information was 'gathered in line with risk to children' (MP5, p44). - One interviewee commented: 'You know people you are talking to and can have informal conversations which can get a lot more done' (MP5, p44). - There was a clearer understanding of 'jargon' used by different agencies. 1.2 Information sharing	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		- Findings suggest that MASH facilitated high quality communication information sharing. - An education welfare officer commented that to begin with there were issues because there was a perception that social services held all information and did not tell MASH professionals because they were not aware they could share information. However, subsequently it was agreed that information could be shared within MASH. 3. The professional in MASH 3.2 Challenges of multi-agency working - Different agencies culture was commented upon by a police officer: 'In the police if you are told to do something, you pretty much do it whereas the social services tend to question a lot more so there were little things … that took a while to become familiar with (M15, p48). 4. The MASH team 4.2 Collegiality and working together - MASH had facilitated an environment where people built positive and effective relationships, and fundamentally inter professional working. This was accredited to co-location and informal case discussions. The strength of team working was reflected by interviewees, especially as professionals felt more supported. 4.3 Culture of MASH - Within MASH, it was seen to develop its own culture. For example 1 interviewee commented: 'this team has a very nice culture of working, everyone gets along, does their job and it's a very nice place to work … professionals working in MASH have been	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		able to meld it together into something good' (MP1, p50). - One police officer commented upon the meshing of police and social work, where they had both 'found a middle ground' (p50) and learnt from each other's styles. 5. MASH and external relationships 5.2 Spreading the word about MASH - A key issue for MASH was the raising awareness amongst other agencies about the role of MASH and the referral process. This was appropriately addressed through outreach by police officers on the street, GPs in practices and local authority training packages. The benefits of raising awareness about thresholds and referral processes were seen which was reflected by the appropriate receipt of referrals to the MASH (MP15, p50). 6. Challenges 6.2 Getting the work done - Staff shortages impacted on workload. - Several interviewees commented on an increase in referrals and services, where staff could not meet the demand. In one instance, an interviewee commented upon heavy workloads by seeing a senior social worker being so busy and working late, which increased stress (MP17, p51). - In contrast, one borough had good resourcing and that meant they could turn around most of the reports within the timescales as risks and dangers were highlighted at the earliest opportunity (MP15, p51). 6.3 IT and technical issues	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		- Many interviewees commented on the multitude of IT systems, notably when an individual professional did not have access to a database they needed or had to travel to a different site to access information from a database that was not available in the MASH.	

6. Mortimer J, North M, Katz A at al. (2012) You have someone to trust - Outstanding safeguarding practice in primary schools. London: Office of Children's Commissioner

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To identify best professional practice in response to child protection and safeguarding concerns in primary schools, including both in-school practice and inter-agency working from the school's perspective. **Methodology:** Qualitative study. Online survey (qualitative), focus groups, semi-structured interviews. **Country:** UK. **Source of funding** Government. Children's Commissioner for England.	**Participants:** Professionals/practitioners - school staff, education professionals and wider school staff and external partners from four schools in a wide geographical area. **Sample characteristics:** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not relevant. **Sample size:** Data presented based on the professionals in educational settings of primary schools: 31 selected school staff 39 education professionals 3 wider school staff and external partners and a range of school staff.	**Narrative findings** Views of educational staff on the challenges of working relationship with other agencies A. Relationships and Communication with other agencies - Maintaining good relationships with outside agencies has enabled schools to trigger appropriate support. 'The more we have engaged in multi-agency working the more we learn how effective it is' (p31). B. Challenges in maintaining relationships with other agencies (p32) 1. Cutbacks were having an impact on provision for vulnerable children through loss of services. 2. Thresholds for some external agencies much higher than schools would like, exacerbated by shrinking funds. 3. Different timescales for other agencies. 4. Lack of understanding of referral protocols to external agencies. 5. Not knowing who to contact. 'We did try to produce a directory of local services/agencies but it keeps changing so hard to keep up. Would be helpful to have one' (member of staff, inner city school, Midlands) (p32). 'Housing - Only now are they starting to understand that they need to be involved with some of our families and attend meetings' (deputy head teacher, inner city school, Midlands) (p32). 6. Tenacity in pursuing relationships with other agencies and findings ways of overcoming these challenges.	**Overall assessment of internal validity:** - Lack of methodological details **Overall assessment of external validity:** + **Overall validity rating:** -

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Follow up: No details on response rates.	'It appears possible to have a high level of neglect/emotional abuse without anyone willing to become involved. In my experience, unless an injury has taken place, no-one is interested' (staff survey, general schools survey data) (p32). C. Key elements for successful multi-agency working identified by staff and external partners (p32) 1. The designated person and other staff are given time to develop relationships. 2. Openness and honesty. 3. Keeping notes and being well prepared for meetings. 4. Mutual respect and good communication. 5. A shared understanding that all agencies were working to the same goal - the best outcome for the child. 6. Commitment and priority given to multi-agency working. 7. Understanding the work of other agencies – e.g. some staff visited the local refuge to try and gain some understanding of the experience of their children (p32). 8. Knowledge and understanding of the people involved (including pupils and parents). 9. Being able to manage stressful situations. 10. Clear boundaries and expectations of each other. 11. Support from the head teacher to deal with issues (p32). 'Often spend hours trying to track down the right person. Just don't give up ... If one agency goes we try another' (special educational needs co-ordinator (SENCO), central London school) (p33).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		D. Practical examples of communication processes (p33) 1. Setting initial planning meetings: with new staff from other agencies to establish protocols, pass on information and clarify purpose and responsibilities. 2. Providing agencies with safeguarding information about the school. 3. Attendance at meetings about children: such as child protection conferences and other multi agency meetings, a high priority. 'More about building relationships with workers and if possible instead of writing letters we will visit them to discuss cases' (pastoral support worker, central London school) (p33). 4. Regular meetings where information is shared by the team within the school: good information sharing, within confidentiality boundaries. 'Referral into our service is well known in the school and is revisited by the SENCO on a regular basis ... the staff raise issues with the SENCO early' (external partner, Central London school) (p33). 5. Providing resources to sustain communication between agencies, e.g., simply providing a venue with the school as a central point for meetings. 'They (the school) have also been good at providing venues for meetings and offering 1:1 support for the child' (external partner, village school, Midlands) (p33). 'The school do not have to pay for the service but they have to donate staff time to work with them and to liaise with other agencies' (external partner, Central London school) (p34). 6. Ensuring a good handover where staff change to include a comprehensive discussion of previous issues, progress made and current work.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		7. Proactively engaging with other agencies, to establish networks through attending meetings, schools felt in a better position to approach relevant agencies when necessary. 'The school knew about the service before it was available here and sought it out and they make good use of it whereas other schools don't' (external partner, Central London school) (p34). 8. Making detailed notes and identifying who is responsible for following up the actions. 'One family has now had in excess of 10 social workers so this has been paramount!' (SENCO, staff survey, outstanding schools survey data) (p34). 9. Having contact details of relevant people to sustain a good level of contact. 10. Persistence for the welfare of the child – following up phone calls and ensuring agreed actions are happening. 'External partners felt supported by the school when chasing up additional agencies 'to ensure action is followed through' (staff survey, outstanding schools survey data) (p34). 11. Reflective practice in partnership working – external agencies be supported by school and be open to discussions if services needed to evolve. 12. Good communication between all sections of the school community: staff, children, parents and other agencies, to provide support strategy at the earliest opportunity, before there was a crisis (p35). E. Relationships and communication with other schools (p37) Excellent sharing of information on vulnerable children a vital component of good practice, particularly in schools with transient populations and at times of	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		transition or transfer, and between the nursery or early years setting and the reception class. This included liaison between the relevant staff and the handover of the records themselves. Good links with other schools prevented children becoming 'lost' to the local authority (p37).	

7. Rouf K, Larkin M, Lowe G (2012) Making decisions about parental mental health: an exploratory study of Community Mental Health Team staff. Child Abuse Review 21: 173–89

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: To explore Community Mental Health Team (CMHT) workers' experiences	**Participants** Professionals/practitioners - staff from the Community Mental Health Team (CMHT) dealing with parents and child welfare/protection: community	**Narrative findings** Practitioners' views on interagency working and decision making (Table 1, p179) 1. The tensions of working across systems	**Overall assessment of internal validity:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
of decisions in the interface between mental health and child welfare. **Methodology:** Qualitative study. Semi-structured interviews and diaries keeping. **Country:** UK. **Source of funding:** Government. Oxfordshire and Buckinghamshire Mental Health Partnership NHS Trust.	psychiatric nurses (CPNs), psychologists, social workers, psychiatrists (mean time since qualification 15.6 years [range 6–23 years], named nurses for child protection (mean time since qualification as health visitors 17.2 years [range 7–24 years]). **Sample characteristics** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socioeconomic position - Not reported. • Type of abuse - Participants were child welfare professionals. • Looked after or adopted status - Not relevant. • Unaccompanied asylum seeking, refugee or trafficked children – Not relevant. **Sample size** n=18 (3 community psychiatric nurses (CPNs), 3 psychologists, 3 social workers, 4 psychiatrists, 5 named nurses for child protection).	a. Felt tensions around their job role, making it difficult to focus on children. This involved issues around role and responsibility, the impact of setting, trying to work systemically, and training and knowledge. 'That's not my job' (Table 1, p179). b. Awareness of power and powerlessness, such as clients as powerless, workers feeling powerless to help. c. Encountering differing thresholds for intervention for children across agencies, e.g., there were inconsistencies in risk thresholds, and unmet needs of children to consider (p179). d. Inter-agency tensions and pressure at work, and the need for prioritising. 'the headless chicken' (Table 1, p179). 2. The dynamics of relationships play in understanding and managing risk (p179). a. Relationship among colleagues, the better the inter- professional relationship, the easier it was to talk about worrying cases. b. Practitioners felt good relationships between colleagues facilitated discussions about stressful cases, and could influence practitioners' risk perceptions, providing reassurance and a place to share ideas, feeling connected, sharing uncertainty and getting reassurance and supported by the team. They needed to feel comfortable about communicating (Table 1, p179).	**Overall assessment of external validity:** ++ **Overall validity rating:** + The methodology was sound but the poor response rate was low.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		c. Team meetings an important place to discuss cases, to share responsibility and to seek reassurance. Some workers did not find team meetings supportive, there is a need to address the culture of team decisions and build cohesion (p186). d. There were mixed experiences of working with social services staff. There were references to delays or failure to act, or 'dumping' by other professionals in the network (p184).	

8. Smeaton E (2013) Running from hate to what you think is love: the relationship between running away and sexual exploitation. Ilford: Barnardo's

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The overall study aims are to: - 'collect data relating to the experiences of young people under the age of 16 who experience both running away and CSE - collect data from practitioners and projects working with young people who experience both running away and CSE - produce an evidence-base that outlines the relationship between running away and CSE and supports recommendations to support policy and	**Participants** Children and young people - 41 young people with experience of running away and CSE whist under 16. Adult survivors of child abuse - 12 of the young people involved were over the age of 18. Professionals/practitioners - 28 projects working with young people experiencing CSE and/or running away via survey; 27 professionals working with young people with experience of running away and CSE via telephone interview. **Sample characteristics** • Age - Young people: 14 n=4 15 n=8 16 n=8 17 n=9 Adult survivors 18 n=7 19 n=2 20 n=2 21 n=1 Professionals - not reported.	**Narrative findings** Already addressed in the findings of Q20: The study reports that the research with professionals and young people identified a range of factors that facilitate and hinder meeting the needs of young people experiencing running away and CSE. 4.1 Resource issues Professionals recommended that there would be more funding available for work with young people running away and experiencing CSE. Professionals identified that practice was facilitated by use of voluntary funds, rather than when money was strictly ring-fenced for particular purposes.	**Overall assessment of external validity** - **Overall assessment of internal validity** + Could be more detailed description of how addressed ethical issues. **Overall score** - Survey of services is entirely of voluntary sector services, and it is unclear whether interviewed professionals represented a wider range of services

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
practice responses to young people who experience both running away and CSE - produce a final report outlining findings, a summary document and a toolkit for practitioners - work with key national agencies to ensure evidence-based findings are incorporated into national policy and practice' (p11). The interviews with young people focused on: '- a history of the young person's life and events and experiences they considered to be important - experiences of running away and CSE - what could have prevented them from experiencing both running away and CSE - their experiences of support seeking - recommendations to both prevent and respond to running away and CSE' (p12). The consultation with professionals focused on:	• Sex - Young people: Female n=25 Male n=15 Transgender n=1 Professionals: Not reported. • Ethnicity - Young people: White British n=32 Mixed Black Caribbean/White British n=3 Mixed Asian/White British n=2 Roma Traveller n=2 Bengali n=1 Sikh n=1 Professionals: Not reported. • Religion/belief - Not reported. • Disability - Young people: Self-defined learning disability or difficulty n=17, this comprised SEN n=9, general learning difficulties n=4, Attention Deficit Hyperactivity Disorder (ADHD) n=2, dyslexia and ADHD n=1, dyspraxia n=1. Professionals: Not reported. • Long term health condition - Not reported. • Sexual orientation - Heterosexual n=29, 'self-defined as gay' (p.13) n=10, bisexual n=1, uncertain about their sexuality n=1. Professionals: Not reported. • Socioeconomic position - Not reported. • Type of abuse - Young people had experienced sexual exploitation. • Looked after or adopted status - Not reported. • Unaccompanied asylum seeking, refugee or trafficked children - Not reported.	4.1.1 Resource issues hindering meeting the needs of young people who experience both running away and CSE Research participants identified the following issues: - Funding cuts, 'predominance of short-term funding cycles' (p59). - Specialised projects cannot meet demand. - Lack of services for young people who experience both running away and CSE, particularly in rural areas. - Lack of appropriate supported accommodation, and use of out of area placements. One professional said: 'We've loaded these children's homes with young people who are at risk of sexual exploitation and it actually destabilise that home so that we can't use it anymore. And then you get young people running away together so they actually strengthen their networks' (p60). Lack of therapeutic accommodation. 4.2 Factors relating to multi-agency approaches to running away and CSE 4.2.1 Factors facilitating general multi-agency working The study identified the following factors: - Effective working relationships with other local voluntary agencies. - Strong relationships with the police. - Effective working relationships with schools. - Working with health professionals and sexual health clinics. - Having health workers based within specialised projects, e.g. having a CAMHS nurse based in a specialist CSE project. - Engagement with A&E departments. - Good relationships with individual social workers. - Having co-located teams with designated workers.	- the voluntary sector perspective of the research is not highlighted or justified in the research methodology. Little consideration in the findings of how contextual and demographic factors shape participant responses.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
the nature of the relationship between running away and CSE '- identification of factors that facilitate projects' work with young people who experience both running away and CSE - identification of factors that hinder projects' work with young people who experience both running away and CSE - identification of groups of young people who experience both running away and CSE that projects find difficult to engage - gaps in national and local policy to meet the needs of young people who experience both running away and CSE - gaps in national and local practice to meet the needs of young people who experience both running away and CSE' (p14). Here we have data extraction information which is relevant to: - young people's views on recognition - aspects	**Sample size** Young people: n=41 Professionals via survey: 28 projects Professionals via telephone interview: n=27.	- Contributing to Multi-Agency Risk Assessment Conferences (MARACs). 4.2.2 Factors hindering multi-agency working - Lack of support from Local Safeguarding Children Boards. - Clash of working cultures between the voluntary and statutory sectors. 4.3 Factors relating to collating and sharing data and information The study reports that professionals noted the importance of collecting and sharing information, particularly when young people move across areas, and that the failure to do this in some places hindered responses to young people who experience running away and CSE. Professionals noted that: - Attitudes to sharing information were important in supporting the work of specialist services. - Agencies may have varying approaches to information sharing. - Specialist projects can be a good source of information. - Missing person reports are also a good way to ensure that children who are experiencing exploitation AND running away are identified. One professional said: 'So when I'm made aware of a young person who may need the CSE service, the first thing I will do is look at the MISPER reports and see how many times that young person has been reported missing. … Because we're getting those MISPER reports on a weekly basis, we can really map out and track someone's progress, someone's deterioration and there's been times when … we've been able to say "well …	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
of professional practice which help/hinder recognition - young people's views on professional responses - aspects of professional practice which help/hinder response to CSE. **Methodology:** Mixed methods. Study comprises: - Interviews with 41 young people who had experienced both running away and CSE whilst under the age of 16. - A survey of 28 projects working with young people experiencing CSE - Telephone interviews with 27 professionals who work with children who experience running away and CSE. The methods section also mentions the development of, and work with, a Research Dissemination Group. However the activities of this		that boy can really do with the X [the CSE] service" and no one's flagging that to us and then I go out to the people involved with that young person, often social care, and say "do you want to refer them to us?". I will say, for example, that not only have they been missing ten times but they have been found at inappropriate addresses, they're found in the red light area, etc. etc. So we can proactively target young people at risk' (p68). 4.4 Professional awareness and knowledge The professional interviewed thought that: - There was a general lack of awareness amongst the statutory sector of running away and CSE, including the perception that running away and CSE is a 'lifestyle choice' (p69). - The concept of 'constrained choice' was useful, which states that 'young people's lack of power relating to age, need and social vulnerability also makes it impossible to give their consent to being sexually exploited' (p69). - Other professionals may be less keen to respond to older children, such as those aged 16 and over. - There is a need for greater awareness about policy and the law in relation to CSE, and of raising general awareness about the issue. - Use of language is important, because young people might not see themselves as having run away they have just 'stopped out' (p71). Young people also emphasised the importance of raising awareness amongst young people and for professionals, through training.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
group are not reported here. **Country:** UK, England. **Source of funding** Voluntary/charity - Comic Relief.		4.5 Factors relating to the local authority Issues not relevant to aspects of professional practice - relates to organisational factors within the local authority. 4.6 Factors relating to the criminal justice system 4.6.1 Young people's experience of being part of a police investigation into CSE Young people who had experienced being part of a police investigation into CSE found this stressful and difficult. This perspective was also emphasised by professionals. 4.6.2 Factors relating to the police Young people: - Emphasised that the police should provide an 'appropriate' (p76) response. One young person said: 'Don't be judgemental when you [police officers] first meet the young person like some police officers when they first met me … [Some] would make a judgement straight away after meeting me' (p.76). - Suggested that the police should give more thought to why they have run away, and that the best option for them may not be to be returned to where they have come from. - Thought the police did not always take appropriate action against perpetrators of CSE, e.g., 1 young person said of an older male perpetrator: 'They never seized his phone or anything and he was grooming me for nine months and I never understood why they [the police] never arrested him … He [the lead police officer on the young person's case] was like "I'm too busy with other cases [against other perpetrators of	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		CSE] you've given me" and I was like "that's no excuse; he could still hurt somebody else [another child or young person]. Why aren't you arresting him?" I got into a massive fight with the police about that and they still didn't do anything so they [the police] just need to take it [sexual exploitation] more seriously' (p77). Professionals: - Noted the importance of the police in responding to CSE, and in having good relationships between specialist CSE projects and the police. However, it was noted that the police response can be variable. Some professionals thought that the police could give 'mixed messages' (p76) about the importance of reporting young people as missing. One professional said 'Parents and carers of young people who are involved in CSE are being told by all professionals about the importance of reporting their child as missing but when they go to the police and they report their child as missing, they're being told that they [the police] are not a taxi service, have you actually looked for the girl – they're only two hours late. You know, they [parents and carers] could be at a child protection conference where someone says 'why didn't you report your child [to the police] whilst she was missing?' Well, I tried to do that but the police told me I was wasting their time'…. Parents can be seen to be not engaging or failing to protect their children when actually they're not but they're getting the wrong response [from the police]' (pp76–7). 4.6.3 Criminalisation of young people who experience running away and CSE	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Professionals noted that young people involved in CSE are often 'criminalised' (p78). One young person also gave an example of this: 'I had a standoff with the police with a knife. For about half an hour I held it [the knife] to myself and was saying that I was gonna hurt myself. They [the police] took it [the knife] off me and arrested me for having an offensive weapon' (p78). 4.7 Factors relating to specialised projects Issues not relevant to aspects of professional practice - relates to organisational factors within voluntary organisations. 4.8 Factors relating to parents and carers of young people who experience running away and CSE Professionals identified the following issues: - The difference that supportive parents and carers can make for young people, and therefore the importance of being able to work with parents and carers, including stressing the importance of reporting young people who run away as missing to the police. - There is a lack of resources to work with parents and carers of young people experiencing running away and CSE. - Professionals need to ensure that they do not seem to be 'blaming' parents and carers for the young people's runaway behaviour. 4.9 Factors relating to direct practice with young people who experience running away and CSE 4.9.1 Factors supporting direct practice The study reports that:	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		Young people identified that: - There is a need for more services 'where they can just turn up' (p82). One young person said: 'There should be places where kids can go to tell someone what's happening to them – someone who will believe them and be able to help them and know what to do for the best. I know there's ChildLine and that but there's some things you don't want to say over the phone. Kids want to go somewhere where there's people they can talk to face-to-face' (p82). - It can be easier to trust workers from the voluntary sector than those from statutory agencies. - It's important for professionals to listen to them. - Outreach work can be valuable. - A good relationship with the worker is 'paramount' (p85), this is supported by an informal manner, and professionals doing what they say they will do. - That it is not always possible to stop young people from running away, but that support should continue to be provided. One young person said: 'I think you have to keep that support in place even when the young person isn't listening and continues to run away and have sex with older men so that when they realise what is going on, there is support in place for the young person … Don't just brush them aside because they can't be bothered at the time. To be honest, it's at the time that they can't be bothered with the support when they don't realise the situation that they're in' (p87). - Ensure a focus on the future as well as the past - Where they had experience of peer support they had found this helpful. One young person said: 'I wasn't taking any notice of what X [the young person's support worker from the specialist sexual exploitation project] was saying so he got this guy who	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		was 18 and had similar experiences as me to come and talk to me … It helped because it was like another view of what I would see in three years' time … He [the 18-year-old male] was like "I know how it is: it's like the best thing in the world and you think they perpetrators] all love you but they do not; they genuinely do not love you; they don't care about you". And it did help me because he had been through this.' (pp88–9). - Young people liked to express themselves using creative outlets. Professionals identified that: - Young people should be able to self-refer to services, and that services should be provided in a 'warm and friendly' environment (p84). - Outreach work can be valuable. - Young people appreciate long term involvement from a consistent worker. - Taking time to build relationship and engage with young people is important. - Flexibility of approach is important, to suit the needs of different young people. 4.9.2 Factors that hinder direct practice Young people identified that: - They would like to see their social worker more often, and felt that statutory practitioners did not always communicate and keep them informed. Professionals identified that: - It is unhelpful to have to stop working with young people when they turn 18.	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		- Some young people are particularly difficult to engage, including: those who are not reported as missing; those who run away for long periods of time; young people who have 'become entrenched in red district culture' (p90); Roma children and young people; 16–18 year olds who have no statutory involvement; young people who have 'fended for themselves' for some time and have become mistrustful of agencies; heterosexual males who are unwilling to disclose exploitation; younger children, such as those aged under 11; young people who have had a lot of professionals involved with them; those who have not developed trusting relationships with any adults; young people with a late diagnosis of ADHD. - Direct work can be hindered by young people's own lack of recognition that they are being sexually exploited. One professional said: 'One of the most difficult ones [hindering factors] is that young people don't recognise their exploitation and so that is a challenge in itself … The young people are needy, they want love and a sense of belonging and that's what exploiters home in on … It may be the only love they've [young people] experienced is in the context of sexual abuse. Some of these children don't have any experiences of safe supportive adults and they distrust professionals, obviously. It's not a quick fix: some of them have really poor attachment history; there needs to be long-term work' (p92). Organisational factor to support projects and practitioners working with young people who run away and experience CSE 1. Lack of organisational support for projects - practitioners felt their work was hindered by a lack of commitment and support from the wider organisation to	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
		ensuring that projects are able to effectively deliver and continue their work, such as the failing to secure continued funding, provide timely response to the increased demand for CSE services and ensure priority of work is shared among other parts of the organisation. 2. Practitioners felt they need appropriate and effective support, quality supervision, training and learning from practice to address and minimise impact upon them; also staff and team professional development, and time to reflect upon their practice '... That is really important in terms of meeting our needs as practitioners [because] it's really challenging, and emotionally challenging, work … Being skilled up and having time to process is really important; having time to think about our work and not having panic knee-jerk responses to things' (p79).	

9. Taylor J, Stalker K, Fry D et al. (2013) Disabled children and child protection in Scotland: investigation into the relationship between professional practice, child protection and disability. London: Scotland. Scottish Government Social Research

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The study reports that 'the aim of this study was to assess how public services (including social work, healthcare, education, police and other related services) identify and support disabled children and young people at risk of significant harm, whether neglect or	**Participants:** Practitioners 'working on issues of disabled children and child protection' (p1). Data was gathered from 21 practitioners across six local authorities via interview, and the remaining 40 through focus groups with five local authority Child Protection Committees. The roles of the practitioners involved are not clear. The research report	**Narrative findings** Practitioners' views on interagency working a. Interagency working was identified as a potential enabler to overcoming lack of individual knowledge and confidence in working with disabled children, in relation to information sharing, crucial in avoiding multiple services repeating interviews with disabled children and for helping to co-ordinate services. There were some failings, tensions and challenges. a. A lack of clarity or context could reduce the usefulness of shared information. In contrast to the comments that each service had a good understanding of	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Limited information on ethical considerations in relation to focus groups.

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
abuse' (p13). The study had the following four research questions: '1. What are the decision-making process and 'triggers' for intervention used by professionals when determining the nature of interventions for disabled children and young people at risk of significant harm? 2. What are specific issues faced by practitioners in Scotland in supporting children and young people at risk of significant harm? 3. How do services co-ordinate to support disabled children and young people at risk of significant harm? 4. What are practice examples in Scotland addressing these issues?' (p13). The findings are not structured according to the four research questions, but according to three over-arching themes which are:	states that 'from each local authority area, potential participants were contacted from social work, education, police, voluntary organisations and health with practice experience of responding to at least two child protection cases involving a disabled child' (p14). However it is unclear who was actually recruited to the study. **Sample characteristics:** • Age - Not reported. • Sex - Not reported. • Ethnicity - Not reported. • Religion/belief - Not reported. • Disability - Not reported. • Long term health condition - Not reported. • Sexual orientation - Not reported. • Socio-economic position - Not reported. • Type o f abuse - Not reported. • Looked after or adopted status - Not reported. • Unaccomnpanied asylum seeking, refugee or trafficked children - Not reported. **Sample size** n=61.	what each other does, there was concern that information shared was not always appropriate (p54). 'Social workers don't know what they're asking for in health, health reports what they think social workers need to know, education also is the same and they'll just say they're fine here or they're not fine here, the environment is different in education …' (p54). b. What is best for the child was mentioned as a focus explaining the ability to keep working together even where relationships were not as strong. 'I think if you've had a poor relationship with another agency you can't carry any grudges you know, you've got to keep working with them for the good of the child' (p55). c. Child protection case conferences, though important for assessing children's needs, were described as unwelcoming, distressing and complicated not just towards children in general, but especially disabled children. The conferences were inadequate when it came to disabled children and young people. '… I think you need to look at whether it's appropriate for the young person to be there or not and whether they understand anything that's going on, and albeit some young people might be twelve or thirteen, they may have the ability of a three year old and I think that needs to be taken into consideration' (p57). 'It's not good at all. Certainly of any of the ones I've been to in the six years I've been here, [children] certainly haven't attended a case conference' (p58). d. Social work often seen by other agencies as having higher thresholds and concerns were expressed by some practitioners that particular children were left in neglectful or risky circumstances for too long.	**Overall validity rating:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
1) The child at the centre; 2) Practice issues (muddling through) 3) Interagency working. We have data extracted findings in relation to theme 1 only, as both 2 and 3 relate more to organisational configuration and training issues. **Methodology:** Mixed methods. Research comprised 'in-depth' interviews with 21 practitioners which included use of a Critical Incident Technique methodology, and 5 focus groups with Child Protection Committees. From the interviews with practitioners, 34 practice examples were developed. The team also developed a series of models to represent the data from the interviews and focus groups. **Country:** UK, Scotland.		e. Health and social services frustrated by the standard of evidence needed by police and courts for criminal prosecutions. Despite the successful adaptations of interviews for children with communication impairments they were still being viewed as unreliable witnesses (p61) or unable to provide the standard of evidence required by the criminal justice system. There appeared to be a difference in the treatment of disabled children compared to non-disabled children and the effects of child protection procedure in practice. f. Current fiscal climate of fewer resources without diminishing demand a potential challenge to disabled children and their families who may require additional support (p75). g. Overall lack of confidence and training among staff when communicating with a child with any communication impairments, being afraid of working on child protection cases involving disabled children, 'muddling through' (p72). h. Safe interagency reflective spaces should be created for discussing and learning from examples of practice related to child protection and disability (p78).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Source of funding: Government. Scottish Government.			

10. Vincent S and Petch A (2012) Audit and Analysis of Significant Case Reviews. Edinburgh: Scottish Government

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Study aim: The aim of the study was to: '… provide key baseline data on the profile, numbers and emerging themes from Significant Case Reviews conducted in Scotland since 2007, and make conclusions and recommendations about the nature and characteristics of factors which can lead to a Significant Case Review, lessons that can be learned both locally and nationally and implications for both policy and practice.' (p30). **Methodology:** Qualitative Study. Analysis of SCR data.	**Participants** Children and young people. The report is based on an analysis of 56 Significant Case Reviews and 43 Initial Case Reviews conducted after 2007. A Significant Case Review is conducted when a child dies and abuse or neglect is identified as a potential factor; if the child or their sibling was on the Child Protection Register (regardless of whether abuse or neglect is suspected as a factor in the death); if the death was accidental or by suicide; if the child was allegedly murdered or died because of a violent act or reckless conduct; or if the child was looked after. Significant Case Reviews are also carried out in cases of significant harm or risk of significant harm as a result of one of the categories of abuse and neglect specified in 'Protecting Children – A Shared Responsibility: Guidance for Inter-Agency Co-operation'. In addition, there must be serious concerns	**Narrative findings** Additional data on interagency working and organisational factors from audits of SCRs (Significant Case Reviews). Common themes: 1. A lack of focus on the child by all agencies, including adult services. A reflective, questioning practice culture be adopted in which practitioners feel confident to challenge parents medical opinion, as well as each other, in order to avoid drift and the operation of the 'rule of optimism'. '… a culture of low expectations and a fatalistic view for some of [the] children' (p64). 2. Managers must listen to frontline staff, acknowledge the difficulties they face in working with troubled families and provide appropriate supervision, training and support. 3. There was confusion in relation to responsibilities in individual cases and a shared understanding of roles across agencies is needed (p79).	**Overall assessment of internal validity:** + **Overall assessment of external validity:** + Information on assessment is part of a broader study. **Overall validity rating:** +

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
Country: UK, Scotland. **Source of funding:** Government. Scottish Government.	regarding professional and service involvement in the case. An Initial Case Review is conducted to determine whether a Significant Case Review should be conducted. **Sample characteristics** • Age - Child - Unborn n=2 (3%); under one year n=21 (30%); 1–4 years n=18 (26%); 5–10 years n=5 (7%); 11–15 years n=19 (27%); 16 years and over n=5 (7%). Mother - Thirty-one reports did not record the age of the child's mother. Where this information was provided, the details were - 20–29 years n=9 reviews; 30–39 years n=13 reviews; 40 and over n=3 reviews. Father - 40 reports did not record the age of the child's father. Where this information was provided, the details were - Under 20 years (17 years) n=1 review; 20–29 years n=6; 30–39 years n=5; 40 and over n=4. • Sex - 13 reviews did not record the gender of the children or young people who were the subject of the review and in 2 cases the child had not yet been born. In those reviews which did provide details on gender 59% (n=33) focused on males and 41% (n=23) focused on females. • Ethnicity - Only 2 reviews recorded details of ethnicity and both children were described as White Scottish. In	4.Thresholds- There was confusion with regard to the status of referrals between different agencies and different professionals. A child protection referral was sometimes regarded to be information sharing or a request for support by the person to whom the 'referral' was made. 5. Because children were considered to be 'in need' as opposed to 'at risk', they were sadly not protected from harm. '… there was a clear failure of all involved services to apply the welfare principle' (p69). 6.Child protection action was significantly delayed due to differences of opinion about which agency should gather information and progress the assessment. 7. Staffing issues Lack of expertise or training in child protection amongst health and social care, including at senior management level (p70). 8. Lack of supervision and support for staff in social work, the police and health visiting. 9. Need for improved single and multi-agency understanding of heavy service demands and better workforce deployment. 10. A multi-agency case management approach at an early stage-assessment of risk to facilitate deci-	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	a number of other cases children and caregivers/families were recorded as speaking languages other than English. • Religion/belief - The report does not state whether reviews included details on religion/beliefs. • Disability - he report states that none '… of the children in this study were recorded as being disabled but the Significant Case Reviews referred to a number of health problems' (p42). The report does not specifically state whether the reviews included details on the disability status of parents or caregivers, however 4 cases appear to have involved a parent or parents with a learning disability. • Long term health condition - The report does not state whether reviews included details on long term health conditions. • Sexual orientation - The report does not state whether reviews included details on sexual orientation. • Socioeconomic position - The report does not specifically state whether reviews included details on socioeconomic status. • Type of abuse - Fatal cases - The deaths of children in these cases were attributed to overdose/drug intoxication n=5 reviews; Sudden Infant Deaths/Sudden unexpected deaths in	sion- making and planning: strengths in practice, particularly in respect of communication, information sharing and responsiveness (p73). 11. Comprehensive or multi-agency assessment of need/risk not always undertaken 12. Concerns regarding communication and inadequate sharing of information within and across agencies, e.g., between forensic CAMHS and GPs; between hospitals and primary care; and between hospitals: 'Systems inhibited the free flow of information particularly between hospitals' (p72). 13. Professionals with important information to share being missing from meetings: the police, the GP, education staff, drug treatment service and criminal justice were missing from an initial case conference. Information not communicated to other agencies or 'became diluted in the translation to attendees' (p75). 14. Cross border communication a challenge when parents moved from a different local authority or a different country. 15. Confusion over roles and responsibilities A lack of recognition of joint responsibility and shared ownership of work with complex families. There was no clear understanding who was 'in charge' of a case (the role of key worker was not always understood by professionals or families leading to confusion as to who was co-ordinating care for the family).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	Infancy n=4 reviews; suicide n=3 reviews; natural causes n=3 reviews; infant sleep related deaths n=3; non accidental injury n=2 reviews; child suffocated after the mother fell asleep during breastfeeding n=1 review; homicide n=2 reviews; death related to bullying n=1 review; unexplained injury n=1 review; fire death n=1 review. The cause was unclear in one review (pending further investigation) and two reviews did not record the cause of death. Non-fatal cases - Physical injury n=11 reviews; ingestion of opiates (i.e. heroin, methadone, etc.) n=6 reviews; neglect n=2 reviews; sexual abuse n=2 reviews; '… concern for unborn child …' n=2 reviews; '… child cruelty and sexual abuse …' n=1 reviews; neglect and sexual abuse n=1 reviews; looked after child convicted of homicide n=1 review; '… safety in care following a complaint by the young person …' n=1 review (p37). • Looked after or adopted status - Nine reviews involved looked after children; and 12 reviews involved children on the Child Protection Register (no further details provided). • Unaccompanied asylum seeking, refugee or trafficked children – The report does not state whether reviews included details on asylum/refugee	16. Procedures- sometimes inadequate. There were not clear pathways and protocols in place for appropriate and timely referral of possible physical abuse cases for specialist investigation and paediatric forensic examination. There were also relationship difficulties between the specialist services in the hospitals which were impacting on patient care' (p76). 17. Recording – not always been fully utilised to inform the planning process. Quality of records problematic, too descriptive and not sufficiently detailed and analytical. There were Inaccuracies and inconsistency in dates. Subsequent information being collected but not added to files. There was evidence of multi-agency planning and reasons for decisions not being clearly recorded (p77).	

Research aims	PICO (population, intervention, comparison, outcomes)	Findings	Overall validity rating
	status or experience or risk of trafficking. **Sample size** 56 Significant Case Reviews and 43 Initial Case Reviews.		