NICE National Institute for Health and Care Excellence

Child abuse and neglect

Economic Appendix C3 – New economic modelling

This report was produced by the Personal Social Research Unit at the London School of Economics and Political Science. PSSRU (LSE) is an independent research unit and is contracted as a partner of the NICE Collaborating Centre for Social Care (NCCSC) to carry out the economic reviews of evidence and analyses.

Authors of the report:

Marija Trachtenberg and Jeni Beecham

Table of contents

C.3.1 Selection of areas for further economic analysis

Table of contents

1 Introduction - economic work as part of guideline development

The economics work is comprised of 2 main components. The first is the critical appraisal and review of existing cost-effectiveness literature and interpreting the results to make recommendations for the UK context. These can be found in Appendix C1 and these are not the focus of this report.

The second component is undertaking new economic analyses. This report, C.3.1, shows the decision-making process and criteria used in deciding which areas should be selected for further economic analysis.

2 NICE guideline methods for selecting areas for new economic analysis

According to the NICE guidelines manual, new economic analyses are beneficial where cost-effectiveness information for an intervention or set of interventions are not available, sufficient, or applicable to the English context. While there are potentially many areas that would benefit from economic analysis, we must prioritise due to time constraints. Criteria for prioritising depends on multiple factors, including the potential net benefit of the analysis, the resource implications of implementing interventions, GC preferences, and the quality and availability of data for economic modelling. This section provides an overview as to how we prioritised areas for new economic modelling.

The prioritisation process first begins with the Economic Plan, which sets out which review questions are particularly important and feasible for analysis. In this case, the Guideline Committee prioritised review question 9 (on early help interventions for children and families at risk of abuse and neglect) and review questions 15 and 16 (on interventions responding to abuse, neglect, or sexual abuse) for further economic analyses given the potential for net benefit and likely data availability. During the guideline systematic review of the research, the Guideline Committee wanted to recommend certain interventions but realised there was a lack of information about their cost-effectiveness. As all of the interventions being recommended had similar resource implications and potential for net benefit, the deciding factors in choosing between them were the quality of the study design and the likelihood of finding data to support economic modelling.

3 Results

Based on the available evidence, a decision was made to undertake further economic analysis in 4 areas:

1. Home visiting for children and families at risk of abuse and neglect (Early help interventions, Review question 9). The results of this analysis are presented in section C.3.2.

   The evidence base for this intervention was based on 1 UK and 11 non-UK studies. However, the initial findings seemed to indicate that

the impact of home visiting was equivocal. However, we thought this might be a result of heterogeneous sample characteristics. Therefore, before any economic analysis was undertaken, tried to determine whether results were less equivocal if the findings were presented for more similar sample characteristics. The findings of that work are presented in section C.3.2. The conclusion of that report is that the impact of home visiting *remains equivocal*, even after presenting results for similar samples. As such, we determined economic analysis was not suitable, and we could not provide additional information on cost-effectiveness. However, the Guideline Committee still decided to keep the recommendation.

2. "KEEP" intervention (Review question 15) – a parenting program for foster carers aimed at improving their parenting skills and reducing child behaviour problems. This intervention forms the basis of Recommendation 1.7.12-13. The evidence for KEEP is based on 1 non-UK study with a large sample size (n=700). The results of this analysis are presented in section C.3.3.

3. "SafeCare" (Review question 15) – a home-visiting intervention aimed at preventing recurrence of neglect among parents with a history involvement in child protective services. This intervention forms the basis of Recommendation 1.7.10. The evidence for SafeCare is based on 1 non-UK study with a large sample size (n=2,175). The results of this analysis are presented in section C.3.4.

4. Trauma-focused cognitive behavioural therapy (T-CBT) (Review question 16) – provided to children and adolescent who have been sexually abused. This intervention forms the basis of recommendation 1.7.17. The evidence for T-CBT was based on 1 good quality meta-analysis (of 10 non-UK studies) with a large combined sample size (n=847). The results of this analysis are presented in section C.3.5.

The conclusion of all three reports (C.3.3, C.3.4, and C.3.5.) is that it is not possible to estimate the interventions' cost-effectiveness given insufficient evidence to make links between the studies' reported outcomes and impacts on wider public sector costs and QALYs. The reports do provide information about the potential resource implications by estimating intervention costs.

Child abuse and neglect

Economic Appendix C3.2

Appropriateness of economic modelling for

Early Help, home visiting interventions

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 (primary prevention)?

This report was produced by the Personal Social Services Research Unit at the London School of Economics and Political Science. PSSRU (LSE) is an independent research unit and is contracted as a partner of the NICE Collaborating Centre for Social Care (NCCSC) to carry out the economic reviews of evidence and analyses.

Authors of the report:

Marija Trachtenberg and Jennifer Beecham

Contents

1 Introduction

Aim

The report looks at whether economic modelling is appropriate for review question 9 on Early Help. It focuses on home visiting interventions for children at risk of abuse and neglect.

The focus is on home visiting interventions, rather than other interventions identified within the Early Help evidence base, because there are more studies of this intervention. The greater the number of studies, the more confident we can be about the precision of the findings.

Economic modelling compares the costs and the effects of two or more interventions. Economic modelling is appropriate when two conditions are met. First, we have information on intervention effectiveness, and second, there is information on resource use (costs). If effectiveness results are equivocal, then economic modelling is not beneficial.

In this report, we describe a more detailed analysis of the evidence. In particular, we explore effectiveness in relation to

- the primary outcome for this question of incidence of abuse and neglect, defined as
    - substantiated cases of abuse and neglect
    - self-reported abuse and neglect or
    - observed measures of abuse and neglect
- the secondary outcome of risk factors for abuse and neglect (for example, depression, stress, family functioning, general wellbeing, attitudes to parenting, etc.).

2 Background

The initial evidence review by the Systematic Review Team indicated that the impact of home visiting on the primary and secondary outcomes was mixed. Another important finding was that the studies' samples were very different. It was thought that the results might be stronger if studies were re-categorised to improve homogeneity.

In particular, we aimed to distinguish between sample members' previous involvement in child protective services, which the literature indicates may be an important factor in intervention effectiveness. We classified studies into three distinct categories: 'primary prevention', 'secondary prevention', and 'mixed prevention'.

- The primary prevention category includes studies where parents had no prior involvement or referral to child protective services.

- The secondary prevention category includes studies where 100% of parents had previous involvement or referral to child protective services.

- The mixed prevention category includes studies in which some of the sample was involved with child protective services but results were reported for the whole sample.

Subsequent sections of this report detail how studies were grouped into these categories.

In this guideline, Early Help is targeted at children who are 'at risk' of abuse and neglect, not those who have experienced abuse and neglect. However, we discovered that some studies' samples had been involved with child protective services before being recruited to the study and therefore may have experienced abuse and neglect. Moreover, not all studies clearly defined what was meant by 'previous involvement in child protective services'. In Section 3.2 we report the studies' own definitions of 'previous involvement in child protective services' but the lack of clarity means that there is some ambiguity in the findings for the secondary and mixed prevention categories.

3 Method

We undertook further analysis on the same studies identified by the Systematic Review Team. We did not extract data on studies identified in the meta-analyses (Elkan 2000; Geeraert 2004, both cited in Barlow 2007) as we wanted to focus on more recent studies. In total, twelve studies were included as relevant to home visiting (Table 1). Five RCTs were identified from two systematic reviews and seven RCTs were identified outside of the systematic reviews. One study is from the UK (Barlow et al.. 2007) and the remaining studies are from the USA.

Further analysis involved extracting data on sample characteristics, where the literature suggests that these may influence intervention effectiveness. This includes child age (including gestation period), whether the intervention was provided in the prenatal or postnatal period, whether mothers were very young, whether mothers were having their first child, presence of domestic violence, and parents' mental health and substance abuse issues. It is worth noting that most studies' participants were mothers even though the interventions were aimed at parents.

We also extracted data on length of follow-up, and primary and secondary outcomes. Outcomes data included the type of outcome as well as the

measurement tool used. For example, the type of outcome might be 'support' and the measurement tool may be the 'Social Support Questionnaire' or the 'Family Resources Scale'. Identifying these items helps us to understand whether it is appropriate to compare studies.

Table 1 Included studies for further data extraction

Identified from systematic reviews (Peacock 2013 or Nelson 2011)			**Identified outside systematic reviews**		
Study	Country	Internal validity	Study	Country	Internal validity
Lowell 2011	USA	Poor (-) / Fair**	Barlow 2007	UK	Moderate (+)
Duggan 2007		Good* / Fair**	DuMont 2008, 2011	USA	Good (++) / Good*
Duggan 2004		Good* / Fair**	LeCroy 2011		Moderate (+)
Bugental 2002		Good* / Fair**	Silovsky 2011		Moderate (+)
Barth 1991		Fair**	Zielinski 2009		Moderate (+)
			Guterman 2013		Moderate (+)
			Green 2014		Poor (-)

Notes
The guideline Systematic Review Team rates studies' internal validity using the following symbols (++/+, or -) where (++) is the highest quality and (-) is the lowest quality. *Peacock (2013) rates this study with a total score of 13-14 out of 15 for study validity (p8). **Nelson (2011) rates this study as having 'fair quality' (pp95–6).

4 Results: definitions, sample characteristics and intervention effectiveness

In Section 4.1 (Table 2), the studies are grouped into the primary, secondary, and mixed prevention categories. In Section 4.2 (Table 3), we then provide the supporting evidence for the new categorisation. In section 4.3, we discuss the similarity of the study samples and the effectiveness findings for both the primary and secondary outcomes. Further details on sample characteristics and findings are in Appendices sections 7 and 0.

4.1 Categorisation of studies

In this analysis of home visiting interventions, 12 studies were included. However, each study can appear in more than one new category if subgroup results are reported separately. This re-categorisation provided a total of 17 sets of results.[1] As Table 2 shows, the primary prevention category has 8 sets of results, the secondary prevention category has 2 sets of results, and the mixed prevention category has 7 sets of results.

Table 2 Study samples grouped into different prevention categories

Primary prevention	Secondary prevention	Mixed prevention
1. DuMont et al 2008, 2011 Subgroup (++) n=179	1. DuMont et al 2008, 2011 Subgroup (++) n=105	1. DuMont et al 2008, 2011 Whole sample (++) n=1173
		2. DuMont et al 2008, 2011 Subgroup: Psychologically vulnerable (++), n=122
2. Lowell et al 2011 Subgroup (-) n=104	2. Lowell et al 2011 Subgroup (–) n=53	3. Silovsky et al 2011 (+) n=105
3. Zielinski et al 2009 Whole sample (+) n=300		4. LeCroy et al 2011 (+) n=195
		5. Barlow et al 2007 (+) n=131
4. Zielinski et al 2009 Subgroup (+) n=unclear		6. Bugental et al 2002 n=96
		7. Barth et al 1991 n=94
5. Duggan et al 2004 n=685		
6. Green et al 2014 (-) n=803		
7. Guterman et al 2013 (+) n=138		
8. Duggan et al 2007 n=179		

[1] Two studies each provided results for two subgroups (Lowell et al 2011; Zielinski et al 2009) and a third study reported results for three subgroups as well as results for the whole sample (DuMont et al 2008, 2011).

Supporting evidence for the new categorisations: defining previous involvement in child protective services

Table 3 (below) shows the studies' own definitions of 'previous involvement in child protective services'. This is important for interpreting the findings and to understand for whom the intervention is effective.

We found that most studies did not clearly define 'previous involvement in child protective services'. Studies used different definitions, which we grouped into one of three definitions. Individuals could have been referred to child protective services (confirmed or unconfirmed abuse), involved with child protective services (and no further detail); or had a history of abuse and neglect.

These categories of definitions indicate a potentially a very broad spectrum of previous involvement with child protective services. Therefore there is a lack of clarity around whether abuse and neglect occurred and the severity of abuse and neglect. In turn, this means it is difficult to interpret findings from studies categorised as secondary and mixed prevention.

Table 3 Definitions of previous involvement in child protective services

Secondary prevention group	
Confirmed report of abuse	
DuMont 2008 (pp13, 14)	• Mothers with previous history, at least one, involved in a confirmed report within 5 years prior to random assignment.
'Involvement' with child protective services	
Lowell 2011 (p197)	• History of child protective service involvement.

Mixed prevention category	
Confirmed or unconfirmed report	
DuMont 2008, 2011 (p304)	• Prior child protective services report, of which 9% had substantiated case of abuse or neglect – over 40% of these prior reports were still open at the time of random assignment to the study.
'Involvement' with child protective services	
Silovsky 2011 (p1437)	• Those with maximum of 2 referrals to child protective services (regardless of substantiation status) were included. However the exact percentage was not provided.
LeCroy 2011 (p1763)	• 24.7% and 11.3% of the intervention and control groups had prior involvement with child protective services.
Bugental 2002 (p246)	• The phrase 'involvement' with child protective services was used but the percentage was unclear. This item was used as a risk factor in the eligibility screening.
Barth 1991 (p365)	• The wording was also 'involvement' with child protective services but the percentage with child protective service involvement was also unclear. However, for 3–5% there was a 'suspicion of previous abuse by mother' from the screen used to determine study eligibility.
'History of abuse'	
Barlow 2007 personal communication	• 'A maximum 20% of the sample had a history of abuse.'

4.2 Effectiveness: primary outcome, incidence of abuse and neglect

4.2.1 Substantiated cases of child abuse and neglect

Main findings

In summary, the effectiveness evidence for reducing substantiated cases of child abuse and neglect remains mixed, and this is true across all prevention categories (Appendix 7). Only half of the samples measuring this outcome found statistically significant reductions.

Primary prevention category

For the primary prevention category, only six of the eight samples measured this outcome. Three samples found statistically significant reductions (Lowell 2011; Zielinski 2009, whole sample and higher risk subgroup) and three samples found no statistical differences (Duggan 2004; Duggan 2007; DuMont 2008). The seventh and eighth samples did not measure this outcome (Green 2014; Guterman 2013). It is difficult to interpret and disentangle the findings because sample characteristics and follow-up periods are different.[2] (Findings are presented in Appendix 7.1)

[2] **Primary prevention category: Studies with statistically significant findings.** Zielinski et al 2009 focused on first time mothers of whom 54% were younger than 19 years old. The intervention was provided during the prenatal period and the study followed the child until age 16 years old. In Lowell et al 2011, mothers' average age was 27 to 28 years old (only 9–10% were younger than 19 years). The intervention was provided when children were between 6 and 36 months old and participants were followed up over a 3-year period.

Studies without statistically significant differences. In Duggan et al 2004 and 2007, mothers mean age was similar, between 23–24 years, of whom approximately 50% were first time mothers. In Duggan 2004, the intervention was provided in the post-natal period, when babies were less than 3 months old. In Duggan 2007, approximately half of mothers were provided the intervention pre-natally, and the remaining 50% when the child was born. In both samples, participants were followed-up until the child was two years old. Both samples had high rates of risk factors for mental health (approximately 50%) and domestic violence (approximately 50%); substance misuse was high in Duggan 2007 (55%) and lower in Duggan 2004 (approximately 21%). In the third study finding no differences, the DuMont et al 2008 sample included mothers who were 19 years old or younger, all of whom were first time mothers. The intervention was delivered during the prenatal period and participants were followed-up until the child was 3 years old. There was insufficient information on the samples' risk factors.

Not measured: Green 2014 and Guterman 2013 did not measure this outcome. Green 2014 included a sample of first time mothers, 11% were younger than 18 years old, 31% younger than 20 years old with a mean age of 22.5 for the whole sample. The intervention was delivered in the post-natal period, when the child was less than 3 months old. The sample had relatively low prevalence of risk factors, with 20% considered having depression, 3% with substance misuse, and 20% having 'relationship trouble'. Guterman 2013 does not provide sufficient information on sample characteristics, the sample includes mothers older than 18, with a mean age of 28 (sd=0.9) and the intervention is provided to children aged 12 years old or younger. Mother's risk factors were unclear (mental health, substance misuse, domestic violence).

Secondary prevention category

In the secondary prevention category, both of the two samples found statistically significant reductions favouring the intervention group. However, we cannot be confident about the consistency and robustness of the findings. This is because of the limited number of studies in this category, which is then compounded by differences in sample characteristics. Samples were different in terms of child age, mother's mean age, and follow-up period and there was no additional information on other sample characteristics, which increases the uncertainty about sample similarity (DuMont et al. 2008, 2011; Lowell et al.. 2011).[3] (Findings are presented in Appendix 7.2)

Mixed prevention category

In the mixed prevention category, most studies reported this outcome (5 of 7 samples) but none found statistically significant differences.[4] As with other categories, these samples were not entirely comparable in relation to child and parent characteristics and time horizon (as detailed below). Therefore, it is difficult to disentangle patterns, if any, between sample characteristics and intervention superiority. (Findings are presented in the Appendix 7.3)

- Timing of the intervention (delivery during the pre or postnatal period) varied. In 2 samples, the intervention was delivered during the prenatal period, in 3 samples, it was delivered in the pre or postnatal period, in 1 sample, it was provided postnatally, and in 2 samples it was unclear.[5]

- Child age and time horizons also varied. One sample included mixed age children, those aged 5 years and younger and measured outcomes over a 1-year period (Silovsky et al.. 2011). Six samples

[3] **Secondary prevention category**: In DuMont et al 2008, 2011, the mother's mean age was 22.5 years old. Exact percentages are not given for whether they were targeted at the pre- (30 weeks) or postnatal period (< 3 months old). Children were followed up until age seven. In Lowell et al 2011, the mother's mean age was 27–28 years old, and children were between ages 6 and 36 months old at randomisation, and were followed up for the next 3 years. There was insufficient information on the proportion who were young, first time mothers, or with mental health or substance abuse needs or were experiencing domestic violence. Socioeconomic status was similar (low, or below high school education).

[4] **Studies measuring this outcome but not finding statistically significant differences**: Silovsky et al 2011; DuMont et al 2008, 2011, whole sample; DuMont et al 2008, 2011 psychologically vulnerable subgroup; Barth et al 1991; Barlow 2007. **Two of the seven samples did not report this outcome**: LeCroy 2011; Bugental et al 2002.

[5] **Mixed prevention category: Gestation period.** Prenatal: Barth et al 1991; Barlow et al 2007. Pre and postnatal period: Bugental 2002, unclear percentage; DuMont et al 2008, 2011, whole sample, 48.5% prenatal at 30 weeks; DuMont et al 2008, 2011, subgroup, psychologically vulnerable, unclear percentage. Unclear: Silovsky et al 2011, children ≤ 5 years old. Post natal: LeCroy et al 2011.

followed children from the prenatal period or shortly after birth. Of these 6, one study followed children until aged 12 months (Barlow et al. 2006), two samples followed children to aged 2 years but these studies did not measure the primary outcome of abuse and neglect (LeCroy et al.. 2011; Bugental et al.. 2002)[6] one sample followed children until they were aged 3 years (Barth et al.. 1991), and finally, two samples had long follow-up periods, measured until the child was aged 7 years old (DuMont 2008, 2011, whole sample; DuMont 2008, 2011, psychologically vulnerable subgroup).

- The proportion of first time mothers was unclear in 5 samples and for the two remaining samples, this percentage ranged between 44% and 54%.[7]

- The percentage of young mothers (<19 years old) was unclear in 5 samples and mixed in two samples.[8] Mother's mean age varied between 22 and 28 years.

- Sample members were generally of low socioeconomic status or completed at or below high school education.

- Mental health needs were present but in unclear proportions in 3 samples. In another sample, mental health needs were present for a small percentage (8–10%), a small to moderate percentage in another sample (35%), approximately 60% in 1 sample, and 100% in 1 sample.[9]

---

[6] Both samples did not measure the primary outcome of abuse and neglect.

[7] **Mixed prevention category: First time mothers**. Unclear percentage: (DuMont et al 2008, 2011, subgroup: psychologically vulnerable; Barlow et al 2007; Silovsky et al 2011; LeCroy et al 2011; Barth et al 1991). Mixed proportion: DuMont et al 2008, 2011, whole sample, 54%; Bugental et al 2002, 44%.

[8] **Mixed prevention category: Young age.** Unclear percentage: Silovsky et al 2011; DuMont et al 2008, 2011, subgroup: psychologically vulnerable; LeCroy et al 2011; Bugental et al 2002; Barth et al 1991. Mixed: DuMont et al 2008, 2011, whole sample, 31% <19 years old; Barlow et al 2007, ±20% <17 years old.

[9] **Mixed prevention category: Mental health needs**. Present but unclear proportion: DuMont et al 2008, 2011, whole sample; Silovsky et al 2011; Bugental et al 2002. Low proportion: Barth et al 1991, 8–10% have history of criminal or mentally ill behavior. Small to moderate proportion: LeCroy et al 2011, p.1765, 35% depressed using CES-D measurement tool. Approximately 60%: Barlow et al 2007. 100% of the sample: DuMont et al 2008, 2011, subgroup: psychologically vulnerable.

- Presence of substance misuse was unclear in 4 samples, present but in unclear proportions in 2 samples, and present for a small percentage in 1 sample (3–15%).[10]

- Presence of domestic violence was unclear in 3 samples, present but in unclear proportions in 2 samples, and low to moderately high percentages in 2 samples.[11]

Conclusions

In conclusion, studies were varied in relation to child and parent characteristics and time horizon. In many studies, important risks and characteristics were not clearly reported. Taken together, they suggest that there is insufficient evidence to determine whether early help interventions are effective (or not) at preventing substantiated cases of child abuse and neglect.

[10] **Mixed prevention category: Substance misuse**. Unclear: DuMont et al 2008, 2011, whole sample; DuMont et al 2008, 2011, subgroup: psychologically vulnerable; LeCroy et al 2011; Barth et al 1991. Present but unclear proportion: Silovsky et al 2011; Bugental 2002. Low percentage: Barlow et al 2007, 10–15% alcohol, 3–10% drugs.

[11] **Mixed prevention category: Domestic violence.** Unclear: DuMont et al 2008, 2011, whole sample; DuMont et al 2008, 2011, subgroup: psychologically vulnerable; LeCroy et al 2011. Present but unclear proportion: Silovsky et al 2011; Bugental 2002. Low to moderately high percentage: Barth 1991, 25–27% of mothers were or are abused; Barlow 2007, 34%.

4.2.2. Self-reported abuse and neglect

This section summarises the number of samples using the same measurement tool to record self-reported abuse and neglect (Table 4). When measurement tools are the same, studies become more comparable. This makes it easier to determine whether interventions are effective in reducing self-reported abuse and neglect.

Comparability and results
Studies rarely used similar measurement tools, and this was true across all prevention categories. Some studies used more than one tool.

Primary prevention category
In the primary prevention category, 5 of the 8 samples measured this outcome. Four samples used the same tool (Revised Parent-Child Conflict Tactics Scale, CTS-PC) but there were only statistically significant reductions favouring the home visiting intervention in three samples, the fourth sample found no differences.[12] However the comparability for this outcome is somewhat limited as one of the samples only focused on some of the subscales (Guterman et al.. 2013). The fifth sample used a different tool, the Adult Adolescent Parenting Inventory, Corporal Punishment Subscale (AAPI-CP), and found no statistical differences (Green 2014).

Secondary prevention category
In the secondary prevention category, neither sample measured this outcome.

[12] **Primary prevention, studies measuring impact on self-reported abuse and neglect:** Dumont et al (2008, 2011) found that, for between ages of 0–3 years old, there were statistically significant reductions for three of six subscales measured: the prevalence (but not frequency) of 'minor physical aggression' in the past year, the frequency (but not prevalence) of 'harsh parenting' in the past week, and finally the prevalence (but not frequency) of 'psychological aggression'. In the child's seventh year, there were statistically significant reductions in the prevalence of psychological aggression and frequency of minor physical aggression (p68).

Duggan et al (2007, p812) used the CTS-PC and found statistically significant reductions in the frequency (but not prevalence) of psychological aggression and mild physical assault in the past year at the child's second birthday. When using the modified version of the CTS, Duggan et al (2007, p812) found statistically significant reductions in the frequency (but not prevalence) of common corporal punishment in the past year.

Duggan et al (2004, p610) reported no differences between groups in areas of psychological aggression, minor physical assault, and severe physical abuse. There were statistically significant reductions for very severe abuse favouring the intervention group. Using the revised measurement tool, the groups were not different on four of five categories, and only having statistically significant reductions in the use of common corporal and verbal punishment.

Guterman (2013, p574) only measured 3 items on the CTS-PC instrument and found no statistical differences for those items over the past 6 months (psychological aggression, physical aggression, neglect).

Mixed prevention category

5 of 7 samples measured self-reported abuse and neglect, of which, some samples used more than 1 measurement tool. The Adult Adolescent Parenting Inventory (AAPI) was used in 2 samples. They were not fully comparable because 1 sample used version 2 of the instrument (LeCroy et al.. 2011). Recognising these limitations, all 3 of the samples found statistically significant reductions favouring the home visiting intervention.

The revised parent-child Conflict Tactics Scale (CTS-PC) was used in 4 samples. This outcome is not fully comparable as each sample adapted the tool either modifying it or by focusing on certain subscales. Given these limitations, three of the four samples found statistically significant reductions favouring the intervention group. However, none of the samples found the same type of abuse or neglect reduced.[13]

[13] **Findings for the mixed prevention group:** 6 of 8 studies measured self-reported abuse and neglect using the same instrument (CTS-PC), but as detailed below each study adapted the tool.

LeCroy et al (2011, p1764) modified the CTS-PC to include only using 'the most serious indicators of abusive and neglectful behavior' and found statistically significant reductions in the prevalence of verbal aggression and minor corporal punishment favouring the intervention group in the child's first year.

DuMont et al (2008, p307), focusing on the whole sample, also used the CTS-PC instrument and found statistically significant reductions mainly in the frequency (but not prevalence of) very serious physical abuse, minor physical aggression, and psychological aggression in the past year and harsh parenting in the past week in the child's first year. In the child's second year, these effects disappeared but there were statistically significant reductions in different subscales: frequency of serious physical abuse and neglect in the past year (p307). The only statistically significant reduction in prevalence occurred in the neglect subscale in the child's first year (p307). In the child's seventh year, there were no differences in prevalence on any subscale but there were statistically significant reductions in frequency of non-violent discipline and serious physical abuse in the past year (DuMont et al 2011, p68).

Silovsky et al (2011, p1441) also used the CTS-PC although excluded one subscale, and statistical power was insufficient for analysis. The study found reductions in non-violent discipline, but the difference on this sub-scale was not statistically significant at follow-up.

Bugental (2002, p251) reported findings for subscales of physical abuse and legally nonabusive use of force (spanking or slapping; pushing, grabbing, or shoving; throwing something at the child), both of which were significantly reduced post-program (1 year, child's first birthday).

Table 4

Findings and comparability of measurement tools for self-reported abuse and neglect

Key:
Bold = statistically significant reduction in self-reported abuse and neglect favouring the intervention group (otherwise no difference)

Self-reported abuse and neglect	**Prevention category**				
	Primary		**Secondary**	**Mixed**	
	5/8 samples		0/2 samples	5/7 samples	
Measurement tool					
Adult Adolescent Parenting Inventory (AAPI)	• **Duggan (2007)** • Green (2014) using Corporal Punishment Subscale			**LeCroy 2011 (version 2)** **Barlow (2006)**	
Revised Parent-Child Conflict Tactics Scale (CTS- PC)	• **Duggan (2007)** • **DuMont (2008, 2011) subgroup** • **Duggan (2004)** • Guterman (2013)	Studies are mostly comparable		**LeCroy (2011)** **Bugental (2002)** **DuMont (2008, 2011) whole sample** Silovsky (2011)	Not fully comparable as each study modified the tool or used different subscales
Self-reported abuse and neglect not measured					
	Lowell (2011) Zielinski (2009) – whole sample Zielinski (2009) – high risk subgroup		Lowell (2011) Dumont (2008) subgroup	Barth (1991) DuMont (2008) psychologically vulnerable subgroup	

4.2.3. Observed measures of abuse and neglect

Observed abuse and neglect was rarely measured so this outcome cannot be used to assess the effectiveness of home help interventions. Moreover, where it was measured, similar issues of non-comparability were found (Table 5).

In the primary prevention category, 2 of 8 samples reported this outcome. One sample reported on the home environment and child's use of A&E and hospital services (Duggan et al 2007) and the other sample measured adequacy of the home environment (Guterman 2013), although using a different measure to that of Duggan et al 2007. In the secondary prevention category, none of the two studies reported this outcome. In the mixed prevention category, 3 of 7 samples measured this outcome. Two samples measured impact on child health using measures of A&E and/or hospital service use (Bugental 2002 and Barth 1991) and the third sample measured the interaction between mother and child (Barlow et al 2007).

Table 5
Comparability of measurement tools for observed abuse and neglect
(Samples can be listed more than once if more than one measure was used).

Observed measures of abuse and neglect	**Prevention category**		
	Primary	Secondary	Mixed
	2/8 samples	0/2 samples	3/7 samples
Measurement tool			
Infant-toddler version of the Home Observation for Measurement of the Environment (HOME) Inventory	• **Duggan 2007** (Favours the intervention, less likely to provide a poor quality home environment)	-	-
Caregiving behaviours Child Well-Being Scales (CWBS) – Household adequacy scale.	• **Guterman 2013 (!!)** (Intervention has significant increase in observed household inadequacy)	-	-
Child Health (including A&E or hospital)	• **Duggan 2007** (No differences)	-	• **Bugental 2002** (Favours intervention) • **Barth 1991** (No differences)
CARE index (mother-infant interaction)	-	-	• **Barlow 2007** (Favours intervention, mothers more sensitive to their babies and babies are more cooperative)
Studies that do not measure observed outcomes of abuse and neglect	• Green 2014 • Lowell 2011 • DuMont 2008, 2011, subgroup • Zielinski 2009, whole sample • Zielinski 2009, subgroup • Duggan 2004	• Lowell 2011 • DuMont 2008, 2011, subgroup	• Silovsky 2011 • LeCroy 2011 • DuMont 2008, 2011, Whole sample • DuMont 2008, 2011, Psych. vulnerable subgroup

4.3. Effectiveness: secondary outcomes, risk factors for abuse and neglect

In relation to the secondary outcome measuring the risk factors for abuse and neglect, both outcome type and measurement tools varied. Table 6 summarises the samples that measured the same type of risk factor while Table 7 summarises the samples using the same measurement tool.

Comparability of risk factors
In the primary prevention category, the most frequently measured risk factors were stress (4/8 samples) and depression (3/8 samples). In the secondary prevention sample, depression, stress, and general mental health or general wellbeing were measured in 1 of 2 samples. In the mixed prevention sample, the most frequently measured outcome was attitudes to parenting (5/7 samples), support (5/7 samples), and depression (4/7 samples). Even where the same risk factors were measured, studies used different measurement tools (Table 7), making comparison across studies difficult.

Table 6
Studies that measured the same type of risk factor
(The most frequently measured outcome is highlighted in bold)

Risk factors for abuse & neglect	Prevention Category			
	Primary (8 samples)	Secondary (2 samples)	Mixed (7 samples)	All samples (17 samples)
Domain				
Depression	3/8	1/2	4/7	8/17
Attitudes to parenting	2/8	0/2	5/7	7/17
Support[14]	1/8	0/2	5/7	6/17
Stress	4/8	1/2	0/7	5/17
Domestic violence	2/8	0/2	2/7	4/17
Substance / alcohol misuse	2/8	0/2	2/7	4/17
General mental health or general wellbeing	1/8	1/2	2/7	4/17
Family functioning	1/8	0/2	0/7	1/17

Total number of studies measuring risk factors for abuse and neglect		
Primary prevention sample	Secondary prevention sample	Mixed prevention sample
• Green 2014 • Guterman 2013 • Lowell 2011, subgroup • Duggan 2007 • Duggan 2004	• Lowell 2011, subgroup	• Silovsky 2011 • LeCroy 2011 • Barlow 2007 • Bugental 2002 • Barth 1991

[14] The measure of "support" was combined to include both measures of internal or external support. An example of internal support would be self-perceived control and social support represents external support.

Table 7
Comparability of measurement tools, risk factors for abuse and neglect

Risk factors	Prevention Category		
	Primary	Secondary	Mixed
Depression	**3 samples**	**1 sample**	**4 samples**
Beck Depression Inventory-2 (BDI-2)			Silovsky 2011
Beck Depression Inventory (BDI)			Bugental 2002
5-item Mental Health Index (MHI-5)	Duggan 2007		
Center for Epidemiological Studies Depression Scale	Duggan 2007 Duggan 2004	Lowell 2011	Barth 1991
Postnatal Depression Scale			Barlow 2007
Pregnancy Risk Assessment Monitoring System	Green 2014		
Attitudes to parenting	**2 samples**	**0 samples**	**5 samples**
Adult-Adolescent Parenting Inventory-2 (AAPI)			LeCroy 2011
Adult-Adolescent Parenting Index (AAPI)	Duggan 2007		Barlow 2007
The Child Abuse Potential Inventory			Silovski 2011 Barth 1991
Mother–Child Neglect Scale (MCNS)	Guterman 2013		
Parent Attribution Test (PAT) – belief about control or power within relationships			Bugental 2002
Graphic depiction of perceived power			
Support	**1 sample**	**0 samples**	**5 samples**
Created for the study			LeCroy 2011
The Emotional/Social Loneliness Inventory			
Adult Hope Scale – Goal setting domain			
Multidimensional Scale of Perceived Social Support (MSPSS)	Guterman 2013		
Pearlin-Schooler Mastery (PSM) scale (sense of control over life circumstances)			Barth 1991
Social Support Questionnaire			Barlow 2007
Social Support: Social Provisions Scale (SPS)			Bugental 2002
Family Resources Scale-revised			Silovsky 2011
Stress	**4 samples**	**1 sample**	**0 samples**
Parenting Stress Index — Short Form	Green 2014 Guterman 2013		
Parenting Stress Index	Lowell 2011	Lowell 2011	
Abidin's Parenting Stress Index (PSI)	Duggan 2007		
Domestic violence	**2 samples**	**0 samples**	**2 samples**
Conflict Tactics Scale (CTS)	Duggan 2004		
Modified Conflict Tactics Scale			LeCroy 2011
Revised Conflict Tactics Scale (CTS2)	Duggan 2007		Silovski 2011
Substance/alcohol misuse	**2 samples**	**0 samples**	**2 samples**
CAGE	Duggan 2004		
Created for the study			LeCroy 2011
Diagnostic Inventory Schedule (DIS) alcohol and drug modules			Silovski 2011
Drug Use Screening Inventory (DUSI)	Guterman 2013		
General mental health or general wellbeing	**1 sample**	**1 sample**	**2 samples**
Brief Symptom Inventory (BSI)	Guterman 2013	Lowell 2011	
State–Trait Anxiety Inventory (STAI)			Bugental 2002
General Health Questionnaire			Barlow 2007
Family functioning	**1 sample**	**0 samples**	**0 samples**
Family Functioning subscale of the Protective Factors Survey	Green 2014		

5 Discussion and conclusions

In terms of triangulating the findings we take a second look at each prevention category and compare findings across the various measures for substantiated abuse and neglect.

In the primary prevention category, similar numbers of samples measured substantiated cases of abuse and neglect (6/8 samples) and self-report (5/8 samples), and fewer samples used observed measures of abuse and neglect (2/8 samples). Greater numbers of samples measured risk factors for abuse and neglect (5/8 samples).

- Three samples found reductions in substantiated cases of abuse and neglect, but none measured self-reported abuse and neglect or observed measures of abuse and neglect. Two samples from the same study found reductions in substantiated cases of abuse and neglect between ages 0 to 16 years old (Zielinski et al.. 2009). The third sample found a reduction in the third and final year of measurement, when children were aged between 3.5 and 6 years old; there were no differences in the two prior years (Lowell et al.. 2011, subgroup).

- Three other samples did not find differences between groups for substantiated cases of abuse and neglect (Duggan 2004; Duggan 2007; DuMont et al.. 2008, 2011, subgroup) but all three samples measured self-report abuse and neglect and all found reductions across different categories. Only 1 sample included observed measures of abuse and neglect, finding favourable outcomes for the intervention. The intervention group was less likely to provide a poor quality home environment, as measured on the Infant-toddler version of the Home Observation for Measurement of the Environment (HOME) Inventory (Duggan et al. 2007).
- Two remaining samples did not measure substantiated cases of abuse and neglect (Green 2014; Guterman 2013) but on the self-report measure, both samples' interventions were not superior to the comparison groups. Only 1 sample included observed measures of abuse and neglect, finding unfavourable outcomes for the intervention as measured on increases in household inadequacy on the Child Well-Being Scales (CWBS) (Guterman et al.. 2013, p574).

The secondary prevention category contained only two samples, both of which found that the early help, home visiting intervention to be superior in reducing substantiated cases of abuse and neglect. One sample found reductions measured between ages 0 and 7 years old (DuMont et al.. 2008) and the other only found reductions in the first year when children were between ages 1.5 and 4 years old, but there were no differences in the

subsequent two years (Lowell 2011, subgroup). Neither study measured self-reported abuse and neglect or observed outcomes of abuse and neglect.

In the mixed prevention category, 5 of 7 samples measured substantiated cases of abuse and neglect and none found the intervention to be superior to the comparison (Barlow 2007; Barth 1991; DuMont 2008, 2011 whole sample; DuMont 2008, psychologically vulnerable subgroup; Silovsky 2011). A majority of samples measuring self-reported abuse and neglect found the intervention to be superior (4 of 5) (Barlow 2007; Bugental 2002; DuMont 2008, 2011 whole sample; LeCroy 2011) the other sample found no differences (Silovsky et al.. 2011). This includes two samples that did not measure substantiated cases of abuse and neglect via reports to child protective services (LeCroy 2011; Bugental 2002). Three samples included observed measures of abuse and neglect. One sample found no differences in child health as measured by use of hospital services (Barth 1991) and one sample found favourable results for the intervention as measured by birth outcomes (Bugental 2002). The third sample found mothers were more sensitive to their babies and that their babies were more cooperative (Barlow 2007).

While we have tried to triangulate and interpret the findings for each prevention category, we are limited in what we can conclude about the effectiveness of early help, home visiting interventions. This is because of the large variation in sample characteristics, time horizons, the limited number of studies on which to draw conclusions, and their heterogeneity. This makes it difficult to understand potential relationships between the sample characteristics and the results, increasing uncertainty about the findings. While we did not extract detailed data on the intervention and comparison group services, these were also observed to be different, adding to the challenge of triangulating and confirming findings.

The aim of this report was to determine whether economic analysis is appropriate given the current evidence base for Early Help, home visiting interventions. In conclusion, an economic model based on either the primary or secondary outcome would not be useful. This is because the effectiveness evidence is equivocal, and therefore insufficiently conclusive for the economic analysis and this is in line ith the findings presented by the systematic review team. This report has shown that there is insufficient information to be certain for whom the intervention(s) is effective, and over what time period it is effective.

6 Additional considerations

In addition to the limitations set out above, other issues to consider are the length of follow-up period, implementation issues, scope of relevant outcomes, and the contexts in which the studies took place.

1) Short and variable follow-up periods

- Most studies followed participants for between 1 and 3 years with only 2 studies measuring impact over a longer time horizon. For 1 study this was from birth until age 7 (DuMont et al 2008, 2011), and for another, from birth until age 16 (Zielinski 2009). In both of these studies the subgroup analyses found different effects.

  Both primary prevention subgroups in the Zielinski (2009) study showed statistically significant reductions in substantiated cases of child abuse and neglect at age 8 and this difference was sustained at 16 years.

  DuMont et al.. (2008, 2011) analysed data for 4 subgroups. Only in the secondary prevention subgroups were there statistically significant reductions in substantiated cases of abuse and neglect. For the primary prevention and the mixed prevention (whole sample) subgroups no statistical differences were found. Results from the fourth subgroup were not reported at 7 years (mixed prevention category, psychologically vulnerable mothers).

- Within the group of studies with shorter follow-up periods (1–3 years) there were 4 samples with mixed ages: 36 months to 3 years old (Lowell et al.. 2011, primary and secondary prevention categories); those aged 5 years or younger (Silovsky et al 2011, mixed prevention); those aged 12 years or younger (Guterman et al 2013, primary prevention). The remaining studies followed children from birth (either pre or postnatally). Across samples, evidence on reducing substantiated cases of abuse and neglect was poor.

2) There may be a need for longer follow-up periods if the intervention is considered to have a ‘lagged’ – or delayed – effect.

- This is important as certain risk factors, such as parent’s mental health, substance abuse, or presence of domestic violence, may require support from additional services. If such services are accessed and are effective in reducing substantiated cases of abuse or neglect, these effects may occur after the observed study period (i.e. ‘lagged effects’).

In our evidence base, where these risk factors were present, there was no information on the proportion of the sample that were referred to services nor the proportion accessing and remaining in contact with services.

3) Intervention implementation

- Implementation issues are frequently cited in the literature and it is important that future research assesses the home visitor's experience of engaging with families and home visitors' referrals to other services. One study found that some home visitors did not feel comfortable talking about the family's risk factors, yet this conversation is key to determining whether additional services are required (Duggan et al 2004, p616; Duggan et al 2007, p819).

4) Scope of outcome measures

- Among studies providing the intervention at the prenatal stage, only a few measured the infant's physical health and associated healthcare service use at birth. This is important because appropriate prenatal care may reduce adverse health outcomes for the child. In 1 study the percentage of mothers with premature births was recorded (fewer for the intervention group), but healthcare utilisation was not measured (DuMont et al 2008, 2011). Likewise, Barth et al. (1991, p368) recorded a range of birth-related outcomes (self-reported) but did not record utilisation of healthcare services.

- A few studies measured impact on the child's general physical health or use of services such as hospital inpatient or A&E admissions. None of the studies measured the parent or child's use of health and social services more broadly; and this is an essential component of economic evaluation.

5) Our report emphasises the need for similarity between samples to help understand potential differences in intervention effectiveness. However, there were additional challenges. We also found that home visiting intervention studies were quite varied. Again, this reduces our ability to compare 'like for like' and therefore understand whether interventions are effective. We found that interventions had different underlying theories of change, ranging from an ecological approach (the family's environment) to a cognitive approach (parent and child behaviours), as well as those that combine these 2 approaches.

6) The impact of this variability is compounded by differences in the study's context such as the threshold for referring to child protective services and thresholds for which cases are substantiated or unsubstantiated. Therefore, it is necessary to triangulate findings by measuring other outcomes, such as self-report and observed outcome measures in addition to measuring health and social care service use.

Implications for practice

Based on this re-analysis of the evidence it is not possible to make recommendations for practice on relative cost-effectiveness.

Implications for policy

This report concludes that there is not enough information to develop an economic model. It does not say that home visiting is ineffective or less cost-effective.

Implications for research

Studies should measure child health outcomes and use of health and social care services. These measures can be triangulated with other outcomes such as substantiated cases of abuse and neglect, self-report, and observed measures of abuse and neglect. It is also important to record home visitors' referrals to services and parents' uptake of services. Future research should also consider using the same measurement tools as used in previous studies (especially where measures are validated) to aid comparability.

Longer follow-up periods may help to capture any delayed effects of interventions on abuse and neglect. Furthermore, longer follow-up periods might be able to capture changes in child health and wellbeing, cognition, behaviour, use of special education services and involvement in criminal activity.

Implementation issues should be recorded, especially in relation to home visitor's experience of discussing and acting on family risk factors.

7 Appendix, sample characteristics

7.1 Primary prevention category

Study info, design, internal validity	Sample size	When the intervention was provided	First time mother	Young age	Age	Socio-economic status (SES)	Mental health needs	Sub-stance misuse	Domestic violence
DuMont et al. 2011, 2008 RCT, ++	Baseline, n=179	**Prenatal** =<30 wks gestation	**Yes (100%)**	**100%** <19	**<19 (all)**	**53%** below high school education	**Small %** (not defined clearly)	**Unclear**	**Unclear**
Zielinski et al. 2009 Whole sample RCT, +	Baseline, I=184, C=116 15-year follow up n=unclear	**Prenatal** <25 weeks	**Yes (100%)**	**47% <19**	**Unclear**	**61%** Low SES unclear how it is defined	**Unclear**	**Unclear**	**Unclear**
Zielinski et al. 2009 Subgroup: high risk (defined as unmarried) RCT, +	Unclear	**Prenatal** <25 weeks	**Yes (100%)**	**47% <19**	**Unclear**	**61%** Low SES unclear how it is defined	**Unclear**	**Unclear**	**Unclear**
Green et al. 2014 RCT, –	n=803, I=402, C=401 Interviewed at 1 year old	**Postnatal** <3m	**Yes (100%)**	**31%** <20, **11%** <18	**Mean** = 22.5	**Unclear**	**20%** depression PHQ-2	**3%** substance misuse	**Unclear**. Relationship trouble= 20%
Duggan et al. 2004 –Primary prevention in relation to index child. RCT, ?	n=685, I=395 C=290 76% eligible participants completed baseline interview and randomisation	**Postnatal** (just after birth)	**Mix. 43-47%**	**NO.** Inferring from 1 SD	**Mean** = 23–24 SD=5.8	**63-67%** Below poverty level.	**43–50%** poor mental health	**19–23%** substance abuse	**43–52%** partner violence
Duggan et al. 2009, 2007, in Peacock et al. 2013 RCT, *?* USA, Alaska	Randomised I=179; C=185; Baseline interview n=325 (90%) I=162 C=163	**Prenatal (41–53%)** Remaining % at child's birth	**Mix. 48-53%** first time mothers	**NO** (inferrring from 1 SD)	**Mean = 23–24** SD=5.7	**57-58%** below poverty level	**52%** depressive symptoms	**55%** substance abuse problems	**45%** yes
Lowell et al. 2011 cited in Nelson 2013 RCT, –	Enrolment, I=78, C=79 6-m follow up I=64, C=67 (82%, 85%) 12-m follow-up I=58, C=59 (74%, 75%).	**Postnatal** 6-36 m	**Unclear**	**Unclear.**	**Mean = 27–28** SD = 7 Range = 17–47	**76-84%** at or Below high school education; "Living in poverty"	**Unclear**	**41–46%** history of substance abuse	**Unclear**
Guterman et al. 2013 RCT, +	Baseline, n=138 I=73, C=65	**Unclear.** **Child is ≤12** years old	**Unclear**	**Mixed, 18+**	**Mean = 28/29** SD=0.9	**Unclear**	**Unclear**	**Unclear %** (receiving treatment)	**Unclear**

7.2 Secondary prevention category

Study info, design, internal validity	Sample size	Child age or gestation period when intervention was delivered	First time mother	Young age	Age	SES	Mental health needs	Substance misuse	Domestic violence
DuMont 2011, 2008 Subgroup: previous CPS involvement RCT, ++ USA	Baseline, n=105	**Unclear.** (Of the whole sample, 48.5% were prenatal at 30 wks and remaining were postnatal, < 3 m)	**Unclear** (Of the whole sample, 54% were 1st time mothers)	**Unclear.** (Of the whole sample, 31% were <19 years old)	**Mean age= 22.5** (SD = 5.5)	**53%** below high school education	**Unclear** .	**Unclear**	**Unclear**
Lowell 2011, in Nelson 2013 RCT, – USA	n=53 subgroup analysis	**Postnatal** 6-36 m	**Unclear**	**Unclear**.	**Mean = 27-28** (SD = 7) Range = 17-47	**76-84%** =< high school education; whole sample “lives in poverty” (unclear definition)	**Unclear**	**Unclear**.	**Unclear**

7.3 Mixed prevention category

Study info, design, internal validity	Sample size	Child age or gestation period when intervention was delivered	First time mother	Young age	Age	SES	Mental health needs	Substance misuse	Domestic violence
DuMont 2011, 2008 Whole sample RCT, ++ USA	Baseline, n=1173 I=579; C =594 Year 1: n=1060 (90%) I=524, C=536 Year 2: n=992, (85%) I=486, C=506 Year 7: N=942 (N=800 kids)	**48.5% prenatal** at 30 weeks Remaining % provided postnatal, <3 months	**54%** first time mothers	**31% <19 years**	**Mean age= 22.5** SD= 5.5	**53%** below high school	**Mixed, unclear %**	**Unclear**	**Unclear**
Subgroup: Psychologicall y vulnerable	n=122	**Unclear**	**Unclear**	**Unclear**	Same as above	**Unclear**	**100%**	**Unclear**	**Unclear**
Barlow 2007 RCT, + UK	RCT, n=131, C=64, I=67	**Prenatal**	**Unclear**	**±20% <17 years**	**Unclear**	**62%** below poverty line	**59- 64%**	**Alcohol** 9.5 -14.9%; **Drugs**, 3% -10%	**34%**
Le Croy 2011 RCT, + USA, Arizona	Baseline interviews n=195, I=98, C=95 6-month interview I=92, C=88 1-year interview I=85, C=86	**Postnatal**	**Unclear**. Average children prior to birth = 2	**Unclear**	**Mean = 23.5–35.4** (p=0.03)	**Unclear**	**35%** depressed using CES-D (p.1765)	**Unclear**	**Unclear**
Barth 1991 RCT, *?* USA	Randomisation: Intervention, N=97 Control, N=94	**Prenatal**	**Mixed**.	**Unclear**	**Mean =** 23.5 years old	Below poverty line & education is below high school	**8-10%** history criminal / mentally ill behaviour	**25-27%** mothers were or are abused	**Unclear**.
Bugental 2002 RCT, *?* USA	Intervention 1: "Unenhanced home visitation", n=31-34 Intervention 2: "Enhanced home visitation", n=32-35 Control: n=25-27	**Mixed.** Late-stage prenatally or soon after birth	**44%** first time mothers	**Unclear**	**Mean age** **Int. 1 =** 25 **Int. 2, =** 26 **Cntrl =** 24	**Unclear**. Average education = 7.5 to 8 years	**Yes. Unclear %.** Item is used as part of screening		
Silovsky 2011 RCT, + USA	n=105, I=48, C=57	**Unclear. ≤5 years old**	**Unclear**		**Mean = 25-28** SD= 7	20% below poverty; 55% ≤ high school	**Included, unclear %.**		

8 Appendix, primary outcome: incidence of abuse and neglect

8.1 Substantiated reports to child protective services

Bold = statistically significant difference between groups (otherwise not significant)

1. Primary prevention category

Primary prevention	Intervention group							Control group							%Difference						
	Cumulative %				Yearly %			Cumulative %				Yearly %			Cumulative %				Yearly %		
	0-2 yrs	0-3 yrs	0-7 yrs	0-16 yrs	1st yr, (1-3.5)	2nd yr (2-4.5)	3rd yr, (3-6)	0-2 yrs	0-3 yrs	0-7 yrs	0-16 yrs	1st yr, (1-3.5)	2nd yr (2-4.5)	3 yr, (3-6)	0-2 yrs	0-3 yrs	0-7 yrs	0-16 yrs	1st yr, (1-3.5)	2nd yr (2-4.5)	3rd yr, (3-6)
Lowell 2011					10%	10%	**15 %**					10%	20%	**30 %**					0%	-10%	**-15%**
Zielinski 2009 subgroup				**19%**							**42%**							**-23%**			
Zielinski 2009 whole sample				**24%**							**32%**							**-8%**			
Duggan 2004	2.1%	3.2%						3.3%	4%						-1.2%	-0.8%					
Duggan 2007	16%							17%							-1%						
DuMont 2008, 2011 subgroup	11%		22%					13%		25%					-2%		-3%				
Green 2014	Not measured (ages 0-1 year old)							Not measured (ages 0-1 year old)							Not measured (ages 0-1 year old)						
Guterman 2013	Not measured, age<12 yrs, 6m follow-up							Not measured, age<12 yrs, 6m follow-up							Not measured, age<12 yrs, 6m follow-up						

2. Secondary prevention category

Secondary prevention	Intervention group				Control group				% Difference			
	Yearly %			Cumulative %	Yearly %			Cumulative %	Yearly %			Cumulative %
	1st yr, (1.5-4 years old)	2nd yr (2.5-5 years old)	3rd yr, (3.5-6 years old)	Age 0-7	1st yr, (1.5-4 years old)	2nd yr (2.5-5 years old)	3rd yr, (3.5-6 years old)	Age 0-7	1st yr, (1.5-4 years old)	2nd yr (2.5-5 years old)	3rd yr, (3.5-6 years old)	Age 0-7
Lowell 2011	**30%**	46%	55%		**55%**	62%	64%		**-25%**	-16%	-9%	
DuMont 2008				**42%**				**60%**				**-18%**

3. Mixed prevention category

Mixed prevention	Cumulative %															
	Intervention group						Control group						% Difference			
	Age 0-1	Age 0-2	Age 0-3	Age 0-7	Age<5 yrs, 1 yr follow-up		Age 0-1	Age 0-2	Age 0-3	Age 0-7	Age<5 yrs, 1 yr follow-up		Age 0-1	Age 0-2	Age 0-3	Age 0-7	Age<5 yrs, 1 yr follow-up
Barlow 2007 UK study	17%						15%						2%				
DuMont 2008, 2011 Whole sample		13%		30%				11%		27%				2%		-10.7%	
Duggan 2007, 2009		16%						17%						-1%			
DuMont 2008 Psych. vulnerable subgroup		ND						ND						ND			
LeCroy 2011		NM						NM						NM			
Bugental 2002		NM						NM						NM			
Barth 1991			ND						ND						ND		
Silovski 2011					20.8%						31.5%						3%

Key: ND = no difference, data not provided; NM = not measured

9 Bibliography

Systematic Reviews

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

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

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

Meta-analyses

Elkan R, Kendrick D, Hewitt M, Robinson JJA, Tolley K, Blair M, et al. (2000). The effectiveness of domiciliary health visiting: a systematic review of international studies and a selective review of the British literature. Health Technology Assessment, 4(13).

Geeraert L, Van den Noortgate W, Grietens H, and Onghena P. (2004). “The Effects of Early Prevention Programs for Families With Young Children At Risk for Physical Child Abuse and Neglect: A Meta-Analysis.” Child Maltreatment. 9(3): 277-291.

RCTs

Barlow J, Davis H, McIntosh E, et al. (2007). Role of home visiting in improving parenting and health in families at risk of abuse and neglect: results of a multicentre randomised controlled trial and economic evaluation. Arch Dis Child. 92: 229–33

Barth RP (1991) An experimental evaluation of in-home child abuse prevention services. Child Abuse Neglect.15:363–75. Journal of Family Psychology, 16(3): 243–58

Bugental D, Ellerson P, Rainey B, Lin EK, Kokotovic A, O’Hara N (2002). A Cognitive Approach to Child Abuse Prevention

Duggan A, McFarlane E, Fuddy L, et al. (2004). Randomized trial of a statewide home visiting program: impact in preventing child abuse and neglect. Child Abuse Neglect, 28: 597–622

Duggan AK, Caldera D, Rodriguez K, Burrell L, Rohde C, Crowne SS (2007). Impact of a statewide home visiting program to prevent child abuse. Child Abuse and Neglect, 31: 829–52

DuMont K, Mitchell-Herzfeld S, Greene R, et al. (2008). Healthy Families New York (HFNY) randomized trial: effects on early child abuse and neglect. Child Abuse and Neglect, 32: 295–315

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

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

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

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

Lowell DI, Paulicin B, Carter AS, Godoy L, Briggs-Gowan MJ (2011). A Randomized Controlled Trial of Child FIRST: A Comprehensive Home-Based Intervention Translating Research Into Early Childhood Practice. Child Development, 82 (1): 193–208

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

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

Child abuse and neglect

Economic Appendix C3.3

Cost-effectiveness analysis

KEEP intervention – Training to improve parenting skills of foster carers with foster children aged between 5 and 12 years old

Review question 15

What is the impact of interventions aiming to respond to children and young people who have experienced abuse and neglect?

This report was produced by the Personal Social Services Research Unit at the London School of Economics and Political Science. PSSRU (LSE) is an independent research unit and is contracted as a partner of the NICE Collaborating Centre for Social Care (NCCSC) to carry out the economic reviews of evidence and analyses.

Authors of the report:

Marija Trachtenberg and Jennifer Beecham

Contents

1 Background and aims

The aim of this report is to undertake a cost-effectiveness analysis based on the KEEP intervention (Keeping Foster Parents Trained and Supported) as described in recommendation 1.6.13.[15] This intervention was selected for further economic analysis due to the lack of cost-effectiveness information suitable to inform UK policy and practice and it was agreed with the Guideline Committee.

The KEEP intervention aims to increase the proportion of positive parenting strategies relative to discipline used (expressed as a ratio) and reduce parent-reported child behaviour problems. This is a 90-minute group-based intervention provided to foster carers on a weekly basis for 16 weeks. Foster carers were paid USD $15 for attending each session. In the evaluation, the KEEP intervention is compared to "care as usual" which are standard caseworker services for foster carers. The effects of the KEEP intervention are measured 5.5 months post-baseline (22 weeks), which is approximately 1 month post-treatment (16 weeks).

The KEEP evaluation was designed as a pragmatic RCT. As it aimed to replicate real world conditions, there were few exclusion criteria. In relation to sample characteristics, the evaluation includes children in both early and late stage of their foster care career (first or multiple foster care placements),[16] and requires that children are in their current placement for a minimum of 90 days[17] and that children are not 'medically fragile'. The intervention is meant to target longer-term foster placements so it excluded children who were in emergency foster care placements and temporary shelters. The majority of foster carers were experienced. Foster carers with their own biological children or were looking after other foster children in their home were also considered eligible and could be included in the study.[18] Foster care placements included both kin and non-kinship arrangements.

This RCT was conducted in the USA with a moderate sized sample (baseline, n=700, follow-up, n=564, 80.6% retention) (Price et al 2008, Chamberlain et al 2008). The study was identified in a good quality systematic review and it was rated as being of moderate quality (Goldman Fraser et al 2013). The systematic review did not provide further detail to explain the rating.[19]

---

[15] The intervention was identified in the systematic search for evidence in review question 15. That question looks at the impact of interventions responding to children and young people who have experienced abuse and neglect.

[16] In this study, children had an average of 3 foster care placements prior to current placement (Chamberlain et al 2008, p.18).

[17] 68% of the sample completed baseline assessments within 6 months of being in their new placement and 76% within 8 months (Chamberlain et al 2006, p.414).

[18] Foster carers had a mean of 2 children in the home (SD = 1) and had an average of 13 previous placements (SD = 4.9) (Chamberlain et al 2006, p.413).

[19] The systematic reviewers rated the evaluation as having medium risk of bias / moderate strength of evidence. The rating indicates that further research *may* change confidence in the

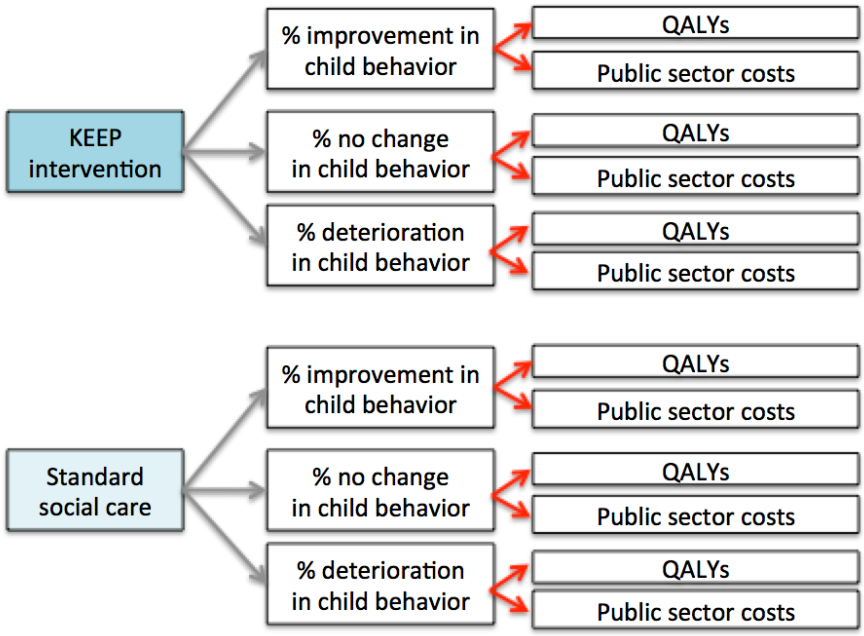
It is important to note that, as this is a US study, we must be cautious about the generalisability of the findings to the English context. For example, if standard social care services in the US generate poorer outcomes than standard social care services in England, then the impact of the intervention in England may not be as large as in the US. Such findings have occurred elsewhere, and while we cannot be certain for this intervention, the issue is worth noting.

2 Methods

Ideally, we would model the reported outcomes from the KEEP intervention (child behaviour problems and positive parenting) to impacts on QALYs and public sector costs. This is illustrated in Figure 1 where we use the example of child behaviour problems although we could have used the example of parenting skills as well.

Given the lack of evidence to make links from reported outcomes to wider impacts on public sector costs and QALYs (as indicated by the red arrows in the diagram in Figure 1), our cost-effectiveness analysis only reports on the additional resource implications (the costs of the intervention) alongside improvements in reported outcomes. The time horizon is the same as in the evaluation: 22-week period for outcomes (5.5 months) and a 16-week period for costs (4 months).

Figure 1 – Modelling structure linking reported outcomes to impacts on QALYs and public sector costs



estimate of the effect and *may* change the estimate (Goldman Fraser et al 2013, p.7). This is in comparison to a rating of low quality, indicating that further research is *likely* to change confidence in findings. The authors assessed quality based on the potential for selection bias, performance bias, attrition bias, detection bias, confounding, and reporting bias (Goldman Fraser et al 2013, p.7). The authors do not provide an itemised list of the evaluation's potential biases, so it is unclear which study design methods were weak.

Our search for evidence linking the study's reported outcomes to wider impacts on public sector costs and QALYs was based on literature searches and liaising with research experts and a sub-group of members from the Guideline Committee. This additional evidence search and results are provided in the discussion section and in the Appendix (Appendix 8.2 - 8.5).

3 Results

3.1 Intervention costs and effects

The US study found that the KEEP intervention led to a reduction in foster carer-reported child behaviour problems and an increase in positive parenting. Both outcomes were measured one month post-treatment (at 22 weeks / 5.5 months post-baseline).

Child behaviour problems were measured using the Parent Daily Report checklist (PDR, Chamberlain & Reid 1987). It asks the foster carer to recall the child's behaviour in the past 24 hours and whether certain behaviours occurred. The measure lists 30 different behaviour problems. Baseline and post-intervention scores were averaged from three telephone interviews (Chamberlain et al 2008, p6).

Foster carer parenting skills were measured as the ratio of positive reinforcement strategies relative to the amount of discipline used. The authors use a 'multi-method index' to compute this ratio based on a list of questions developed by the study authors.

Effectiveness results
The KEEP intervention led to a mean reduction of 1.22 foster carer-reported child behaviour problems and an increase of 0.07 in the ratio of positive reinforcement strategies relative to the amount of discipline used (Table 1).

Table 1 – Effectiveness results

Outcome measure 5.5 months post-baseline	**Whole sample**	
	Intervention, n=359	Control, n=341
Mean number of child behaviour problems as reported by the foster carer	A mean reduction of 1.22 child behaviour problems as reported by the foster carer, favouring the intervention group. Standard error[20]=1.66 Small effect size, Cohen's d=0.26	
	Baseline 5.92 (4.26) Termination 4.37 (3.91)	Baseline 5.77 (3.93) Termination 5.44 (4.15)
Mean change in the proportion of positive parenting relative to the amount of discipline used by the foster carer	A mean improvement of 0.07 in the proportion of positive reinforcement relative to the amount of discipline used by the foster carer, favouring the intervention group. Standard error[21]=0.004 Small effect size, Cohen's d=0.29	
	Baseline 0.53 (0.27) Termination 0.60 (0.28)	Baseline 0.52 (0.27) Termination 0.52 (0.28)

Costs
We estimated the UK-equivalent costs of the KEEP intervention using a *full-cost approach*, which is considered a best practice standard in economic evaluations. *A full-cost approach includes* wages, oncosts (pension and

[20] The standard error was calculated from the study using the information provided. First, the pooled standard deviation of the differences between the 2 groups was calculated by rearranging the equation for effect size (ES=difference in group means between time 1 and time 2/pooled standard deviation). The available information (effect size, 0.26, mean difference, 1.22) gave a pooled standard deviation=22.02. Second, the standard error was calculated using the following equation: $\sqrt{((sd^2/n_a) + (sd^2/n_b))}$ where sd^2 is the pooled standard deviation (22.02) and n_a and n_b are the sample sizes of the two groups. The standard error of the difference between two groups is 0.166.

[21] Calculated using the method above.

national insurance contribution), training, qualifications, direct and indirect overheads, capital costs, and the time cost of travel. Intervention costs reflect 2015/16 prices, based on the Unit Costs of Health and Social Care compendium (Curtis and Burns 2015). The costs of the intervention were estimated based on the information provided in the study, which is described subsequently. As will be noted, there were several factors that influence the intervention cost estimates. For this reason, we provide a range of cost estimates, indicating the lower- and upper-bound intervention costs.

The two main factors that influence intervention cost are group size and the type of professional delivering the intervention. The authors report that the KEEP intervention was delivered in groups of 3 to 10 foster carers led by a trained facilitator and co-facilitator who are both para-professionals (the authors do not provide any further information as to the facilitators' qualifications and professional background). Given the potential for costs to vary depending on group size and type of professional delivering the intervention, we provide cost estimates for group sizes of 3, 6, and 10. Furthermore, as the authors do not elaborate on the professional background of the facilitators, we estimated a lower and upper cost estimate by assuming that a family support worker or a child social worker could deliver the intervention.

Other intervention costs include the costs of recruiting foster carers and the cost of additional intervention components such as provision of childcare, refreshments, venue hire, and materials. The facilitators' time-cost of travelling to the sessions are also included, which we assumed to be 1 hour of travel for each session, for each facilitator. We also included the costs of repeating the session with a home visit when sessions were missed, which occurred in 20% of total sessions (Price et al 2008, p.6). We also included the costs of facilitator training, which lasted for 5 days, and the costs of supervision, provided to review and discuss videotaped sessions (Price et al 2008, p.6).

Intervention costs also include payments to foster carers for each session attended. In the US study, foster carers were paid USD $15 per 90-minute session attended ($10/hour). This represents an hourly payment that is approximately 1.5 times the minimum wage (varies state to state). Our costing approach estimates the English-equivalent payment to be £10.80/hour (1.5 x national minimum wage of £7.20/hour) and therefore a payment of £16.20 per 90-minute session (1.5 x £10.80/hour).

Appendix 8.1 provides the details of the intervention costing.

Taken together, *and using a full economic costing approach*, intervention cost estimates range from a low of £2,000 per foster carer (delivered by a family support worker for a group size of ten) and can be as high as £9,800 per

foster carer (delivered by a child social worker for a group size of three) (rounded to the nearest hundred) (Table 2).[22]

Table 2 - Total cost per foster carer, considering different group sizes and depending on the type of professional delivering the intervention

Scenario	Group size 10	Group size 6	Group size 3
Family support worker	£2,012	£3,100	£6,166
Child social worker	£3,121	£5,000	£9,818

3.2 Cost-effectiveness analysis: how the results were calculated

Our results are calculated and presented in 2 different but complementary ways, both of which help in deciding whether the KEEP intervention is cost-effective.

In the first approach, we present the findings as a mean cost-effectiveness ratio. This is defined as the mean cost for a 1-unit improvement in the outcome. For the results of this intervention, this is presented as the mean cost for an improvement in positive parenting strategies relative to the amount of discipline used and the mean cost for a 1-unit reduction in child behaviour problems (Table 3).

The cost-effectiveness ratio is useful when you must select between different interventions that measure the same outcome. A decision maker would want to choose the intervention that has the lowest cost per unit of improvement.

A decision-maker may also have a limit as to how much they are willing to spend for an additional unit of effect. This is termed the 'cost-effectiveness threshold'.

The only outcome for which there is an established threshold is the quality-adjusted life-year (QALY). The QALY is a measure of health-related quality of life. NICE guidance suggests that interventions with cost-effectiveness ratios equal to or less than £20,000 to £30,000 per QALY are cost-effective, in areas where the QALY is applicable. However, for non-QALY outcomes, there is no

[22] Our cost estimate for an intervention delivered by a family support worker for a group size of 10 is higher (£1,800 per person) than estimates produced for other group-based parenting interventions provided in the PSSRU Unit Cost volume (Curtis and Burns 2015, p.99). The median and mean cost estimates per person for other group-based parenting interventions are £1,005 and £969 respectively (Curtis and Burns 2015, p.99). However, our intervention cost estimates doubles if the intervention is delivered by a child social worker (£2,900 per person) (and group size remains at 10). Our intervention costs may be higher because the duration of the intervention was longer (16 vs. 12 weeks) and because we included the costs of recruitment.

guidance around an acceptable cost-effectiveness ratio.[23] This means there are no established thresholds for determining the cost-effectiveness of an intervention based on the outcomes of child behaviour problems and positive parenting, unless there was some way of linking this to QALYs. However, as we mentioned earlier in the section on methods, we were unable to find evidence that makes these links.

Had we been able, we could compare the KEEP intervention to any other intervention and any other population so long as results are measured with QALYs. For example, a decision maker could decide between the KEEP intervention and an intervention for abusive parents.

In the second approach, we present the results using a cost-effectiveness acceptability curve (CEAC) (Figure 1 and Figure 2). The CEAC summarises the parameter uncertainty surrounding the intervention's cost-effectiveness ratio. Parameter uncertainty relates to the variation in effectiveness, which considers the distribution of effect, rather than using only information about the mean effect. With that information, it is possible to provide a probabilistic estimate of the intervention's cost-effectiveness. We undertook a probabilistic analysis using a simulation, in particular, a Monte Carlo simulation. In our analysis, we only simulated outcomes, and not costs. We do not use probabilistic sensitivity analysis for costs because intervention inputs are fixed, and not probabilistic. As there is a range factors influencing intervention costs, it is more appropriate to use scenario analyses, as presented earlier. Instead of performing six scenario analyses, we use only the lower and upper intervention cost estimates (£2,012 and £9,818 per foster carer).

The probability of the intervention's cost-effectiveness is calculated in several steps. First, we create a hypothetically large sample size (we chose 1,000, as this is standard practice). We then created a randomisation formula that picks a value, at random, within the constraints of the mean and standard deviation for the outcome. In the absence of information about the distribution of positive parenting and distribution of child behaviour problems, we assumed that the chances of getting different numbers are based on a 'normal' distribution. The 'normal' distribution assumes that chances of getting different numbers on either side of the mean is symmetrical, and that numbers closer to the mean have a higher chance of occurring than values farther away from the mean. Effectively, we are assuming that a majority of the sample will have an effectiveness outcome that is close to the mean score, and a smaller percentage of the sample will have effects that are either smaller or larger than the mean.

When the probabilistic analysis is set up, it is able to tell us the probability that the intervention is cost-effective at a certain amount of money. For example, a decision maker may only be willing to pay £20,000 per QALY. The CEAC tells

[23] QALYs only measure health-related quality of life. While it is useful, we must recognise that, for social care interventions, there are other non-health related outcomes that are equally important. At this point in time, NICE has not yet determined the social-care equivalent QALY.

you the probability that the intervention is cost-effective at various amounts that a decision-maker is willing to pay. This is calculated by finding the percentage of times that the intervention is cost-effective at or below £20,000, and this is done over a large number of simulations (in our case, 1,000 simulations). For example, we might find that 75% of the 1,000 simulations result in a cost-effectiveness ratio equal to or below £20,000. However, as discussed earlier, when we have cost-effectiveness ratios with non-QALY outcomes, there is no established threshold for assessing cost-effectiveness. In any case, presenting results using a CEAC is still helpful because it summarises the uncertainty surrounding the intervention's cost-effectiveness ratio.

3.3 Findings from the probabilistic analysis (Monte Carlo simulation)

3.3.1 The mean probabilistic cost-effectiveness of the KEEP intervention

The KEEP intervention was able to produce a mean reduction of 1.22 child behaviour problems (standard error=1.66). When we use this data in combination with intervention cost estimates in a Monte Carlo simulation, this corresponds to a mean probabilistic cost-effectiveness ratio between £1,583 and £8,429 per one-unit reduction in child behaviour problems (representing the lower and upper intervention cost estimates, respectively) (Table 3).

Likewise, the KEEP intervention was able to improve the foster carers' parenting skills (expressed as a ratio) by a mean of 0.07 (standard error = 0.004). This corresponds to a mean probabilistic cost-effectiveness ratio between £28,777 and £140,379 per one-unit improvement in the ratio of positive reinforcement relative to the amount of discipline used (representing the lower and upper intervention cost estimates, respectively) (Table 3).

Table 3 – Mean probabilistic cost-effectiveness ratios

Intervention cost scenarios		
Lower	Upper	Outcomes
£1,583	£8,429	One-unit reduction in child behaviour problems
£28,777	£140,379	One-unit improvement in parenting skills

3.3.2 Cost-effectiveness acceptability curves (CEAC)

In relation to the CEAC for parenting outcomes (Figure 2), the simulated data show that there is a 100% probability that the additional cost of the KEEP intervention, compared to standard social care services in the US study, is less than £34,500 for a one-unit improvement in positive parenting relative to the amount of discipline used (in the lower intervention cost scenario). In the higher cost scenario, there is a 100% probability that the additional cost of the KEEP intervention, compared to standard social care services in the US

study, is less than £163,500. These estimates are based on the simulation of the effectiveness data, which find a 100% chance that the intervention will result in a positive impact on parenting skills, based on the findings of the US study (Figure 3).

Figure 2 – CEAC: positive parenting

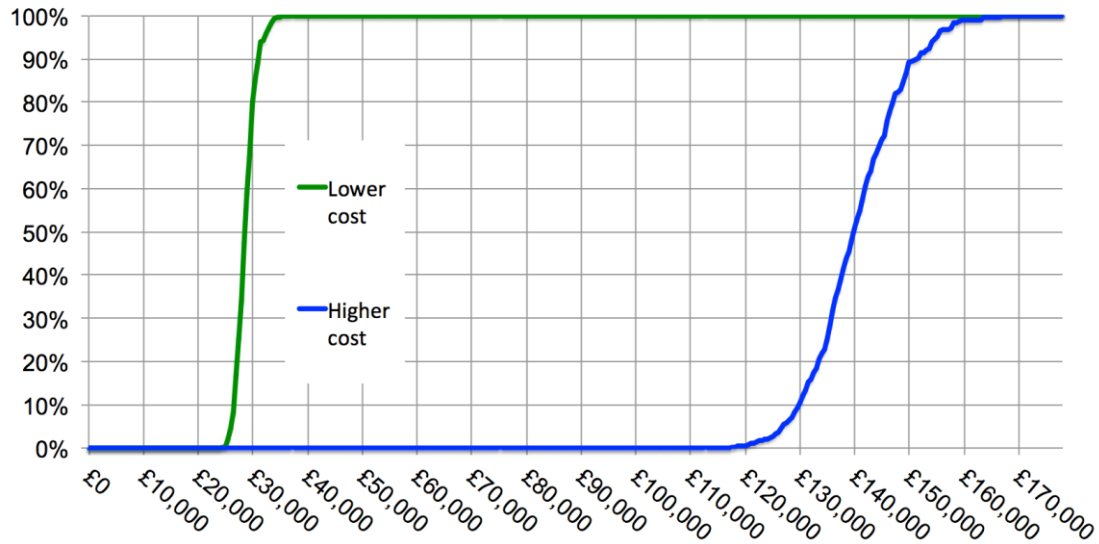
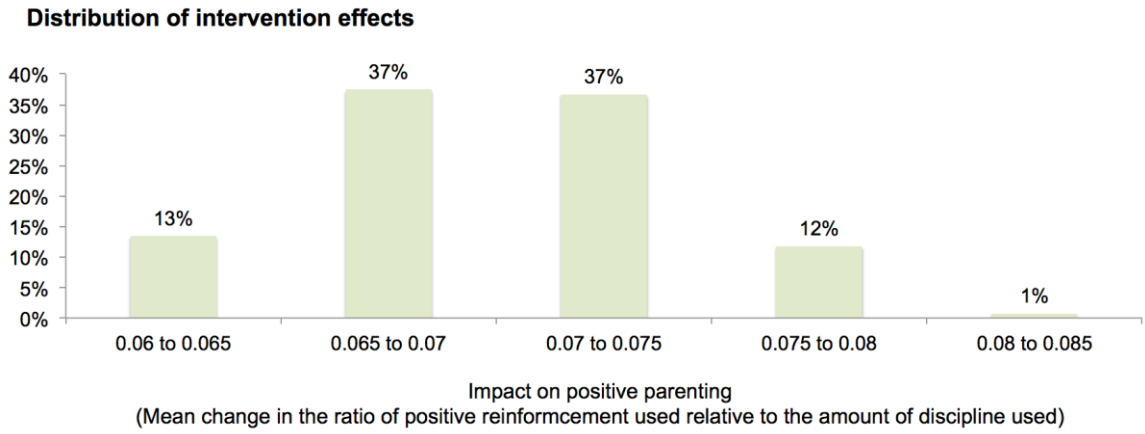


Figure 3 – Probabilistic simulation: distribution of effects on positive parenting skills



In relation to the CEAC for the outcome of child behaviour problems (Figure 4), the data show that there is a 70% probability that the additional cost of the KEEP intervention, compared to standard social care services in the US study, is less than £6,000 per one-unit reduction in child behaviour problems (in the lower-cost scenario). In the higher cost scenario, there is a 70% probability that the additional cost of the KEEP intervention, compared to standard social care services in the US study, is less than £31,000 per one-unit reduction in child behaviour problems. These estimates are based on the simulation of the effectiveness data, which find a 77% chance that the

intervention will result in a positive impact on child behaviour problems, based on the findings of the US study (Figure 5).

Figure 4 – CEAC: child behaviour problems

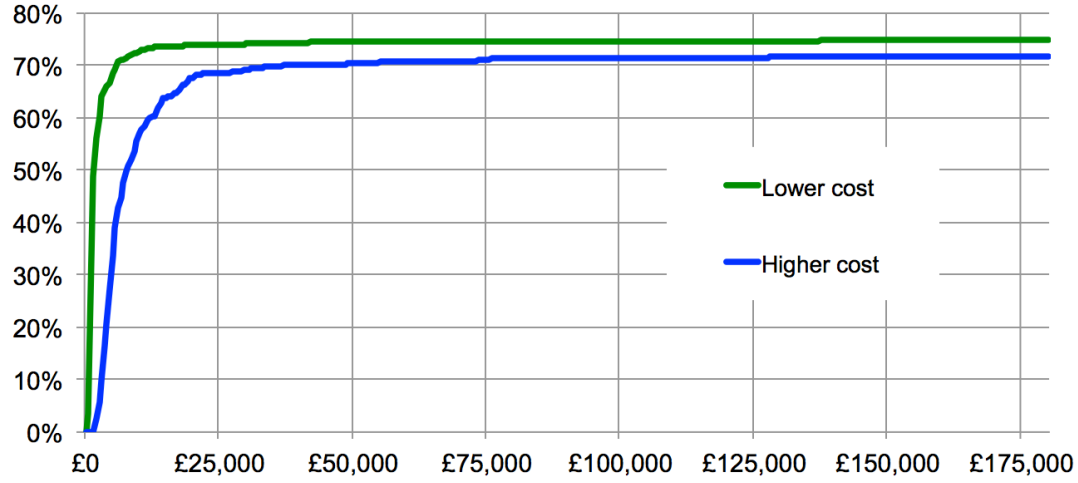
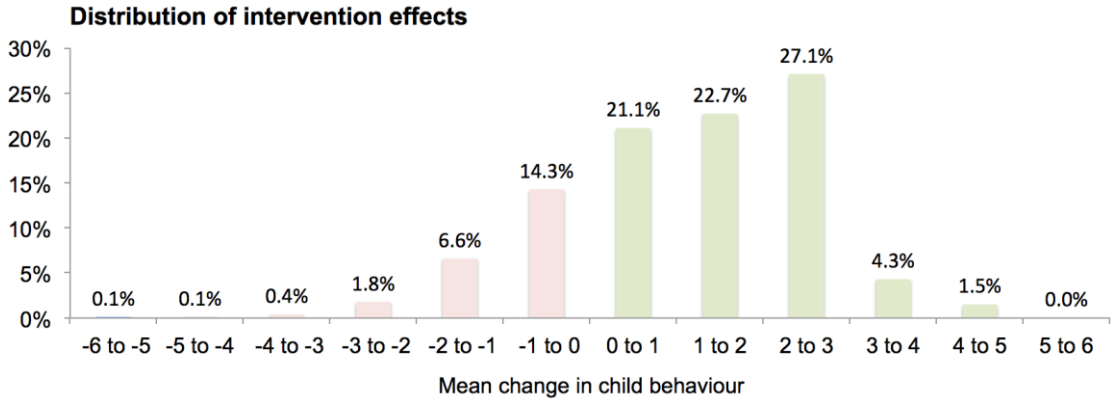


Figure 5 – Probabilistic simulation: distribution of effects on child behaviour problems



4 Discussion

Findings

The US study found that the KEEP intervention led to a mean 0.07 improvement in the ratio of the parent's use of positive reinforcement relative to the amount of discipline used. The KEEP intervention also led to a mean reduction of 1.22 foster carer-reported child behaviour problems per day (measured one-month post intervention, 5.5 months (22 weeks) post-baseline).

The costs of the intervention calculated here were lowest if delivered by a family support worker to a group of ten foster carers (£2,012 per foster carer). Intervention costs per foster carer were highest if delivered by a child social worker to a group of three foster carers (£9,818 per foster carer). These costs are inclusive of payments to foster carers for attending each session

(estimated £210.60 per foster carer (or £16.20/session) assuming that foster carers attend 13 of 16 sessions (as it was reported that foster carers missed 20% of sessions). It is very important to note that this does not imply that the lower-cost intervention is more or less cost-effective than the higher cost intervention. The US study was not designed to answer this question, as it did not compare the effectiveness of KEEP based on differences in professionals' qualifications and background. The US study only reports that the KEEP intervention was delivered by trained para-professionals, without further elaboration as to their background. The use of para-professionals suggests that less senior staff (and less costly staff) could deliver the intervention with appropriate training. However, given the absence of information on para-professionals' seniority, we simply provide a potential range of intervention cost estimates – a lower and upper bound estimate – which takes into account lower and higher levels of seniority.

Strengths and limitations

Our analysis is limited to estimating the resource implications of the intervention and reporting the study's findings on foster carer parenting skills and child behaviour problems. We did not identify robust evidence to make links to QALYS and public sector service use based on the reported outcomes in the US study. However, the strength of this analysis is that we provide a narrative on the potential downstream consequences. These were derived from additional literature searches and through consultation with research experts and members of the Guideline Committee. From this process we identified potential benefits and cost-savings in the short, medium, and long-term. These additional benefits and cost savings are discussed below.

Potential benefits and cost savings

In the short term, some of the intervention costs might be offset by preventing a foster placement disruption (change in foster carers) or the costs of a child running away. The US study found that children with 7 or more foster carer-reported child behaviour problems per day had a higher chance of a placement disruption. In particular, each additional behaviour problem above 6 per day had an additional 1.2x higher chance of a placement disruption (Chamberlain 2006). For example, a child with 10 foster carer-reported child behaviour problems per day would have 2.07x higher chance of a placement disruption than a child with 6 and fewer behaviour problems (calculated as 1.2^4). However, we do not know the distribution of the number of child behaviour problems among foster children. If there are a large proportion of foster children with higher numbers of child behaviour problems, there is greater scope for cost-savings. However, if there is a smaller proportion, the scope for cost-savings may be smaller. However, we were unable to find data on the distribution of child behaviour problems among foster children when using the Parent Daily Report checklist measurement instrument.[24]

Evidence from this US study is also supported by evidence from one meta-analysis of unclear quality (Oosterman et al 2007, p.67). The results from the

[24] Nor were we able to find data to crosswalk from the Parent Daily Report checklist to a more widely used measure, such as the Child Behaviour Checklist.

meta-analysis are limited in some important ways. Most importantly, only two studies undertook multivariate analyses while the other 11 undertook univariate analyses, which may overestimate effects of association. Furthermore, due to a lack of reporting, it is unclear what definition was used for foster placement disruption, and therefore, whether it is applicable to our analysis. In light of the limitations, the meta-analysis finds an association between child behaviour and foster placement disruption, which range from small (r=0.22 to 0.28) to moderate (r=0.51). The size of the effect depends on whether univariate or multivariate statistical analyses were used, with multivariate analyses showing smaller effects. These findings are also supported by another systematic review, but these results are summarised narratively and do not use meta-analytic methods (Rock et al 2013).

This finding is important because foster placement disruptions can be costly. We did not find studies that calculated the societal or public sector costs of a foster placement disruption. However we did identify the administrative costs to children's social care services to be between £2,150 and £2,700 per change in foster placement, depending on whether the child is placed through the local authority or using an independent foster care agency (Curtis 2015, p.94 and p.130).[25] These costs assume that a new placement requires social care processes 2 through 6, which involves care planning, leaving care, finding a new placement, and a review of the new placement. Furthermore, costs might be higher if the initial disruption requires an emergency placement followed by a long-term placement. However, we could not find data on the probability of a 'smooth' disruption versus a more challenging disruption requiring two placements. Furthermore, disruptions are likely to have administrative costs to health, social care, and education services if subsequent placements were made out-of-area (Ward et al 2009, p.1117)[26]. National data indicate that 37% of new placements are made outside the council boundary and 5% are made 20+ miles away from inside council boundary (Department for Education, 2016, Tab A7, B3).

In the medium-to-long-term, the chances of having a future foster placement disruption are increased when a child has a history of frequent foster placement disruptions. Evidence from the same meta-analysis (above) finds that the number of previous out-of-home placements is associated with an increased likelihood of future foster placement disruptions (small effect, r=0.12, based on 5 studies, p<0.001) (Oosterman et al 2007, p.66). Indeed, the US study found that the KEEP intervention had a preventive effect which reduced the risk of foster placement disruptions among a subgroup of children who had a high number of prior placements. In particular, the KEEP intervention reduced the risk of a placement disruption by 12% (Price et al 2008, Table 3, p.18).[27] This means that there are potential cost savings in the

[25] Total costs based on: process 2, care planning, process 4, leaving care, process 5, finding a new placement, and process 6, review.
[26] Placements made outside of council boundaries may result in additional administrative costs due to changes in education (schools and potentially special education assistants) and healthcare providers (GPs, dentists, mental health or occupational therapists).
[27] Standard error = 0.06, p=0.05.

medium-to-long term, as the KEEP intervention reduces the risk of future placement disruptions among those with a history of placement moves.

One research expert believes that the largest cost savings would occur in the long-term if the intervention leads to foster children becoming more settled in long-term foster care and continue to stay through their adolescence (Ian Sinclair, personal communication, April 2016). This is opposed to a trajectory where the child has difficulties with their foster placement and eventually moves into residential care, which is more costly than foster care services. The cost of residential homes for children is between £2,900 and £3,170 per week, which is 4 to 5 times higher than the cost of foster care placement, at £614 per week (Curtis 2015, pp. 84-86).

In relation to additional effects on the child, the literature suggested the potential for a positive effect on the child wellbeing as a result of preventing a foster placement move (see Appendix 8.2 for more detail). Our search of the literature did not identify any English studies but we did find three US studies finding a *causal* link between foster placement moves and a subsequent *negative* impact on child wellbeing. These three US studies are important because they were designed to assess causality and so this evidence is stronger than the many studies available that only measure *association* (Aarons et al 2010, Rubin et al 2007, Newton et al 2000). In each study, child wellbeing was measured using the Child Behaviour Checklist (CBCL). The CBCL is composed of three components: internalising and externalising behaviour problems and total behaviour problems. The three US studies show consistent results. After controlling for initial baseline characteristics, including child behaviour, placement instability or a high number of placement changes, placement moves result in the child having subsequently greater number of behaviour problems. The magnitude of the effect reported in each study varies depending on: sample characteristics, type of data used, selection of covariates, time horizon, and statistical methods. The limitation in this evidence base is that these studies measure foster placement moves, and not foster placement disruptions. Foster placement moves could be made for either positive or negative reasons whereas a foster placement disruption is more likely to be associated with challenging, and potentially more negative circumstances. Therefore the results of the meta-analysis are not entirely applicable.[28]

In summary, the additional research suggests the potential for reduced costs to the public sector in the short, medium, and long-term and benefits to child wellbeing as a result of the interaction between reduced child behaviour problems and foster placement disruption. However, we must emphasize that without a robust economic evaluation, we cannot be sure whether or not the intervention is cost-effective, based on currently available evidence.

Barriers to undertaking economic modelling for the English context

[28] It is worth noting that the authors of one study (Aarons et al 2010) recommend more research in this area to do just that. They also suggest further research that covers a time horizon longer than 3.5 years and investigate the impact of age and gender as moderating factors.

As described earlier, the US study found that the KEEP intervention had a greater effect among children with initially higher numbers of behaviour problems (Chamberlain et al 2006, 2008). We thought it would be worth finding out the differences in the intervention's cost-effectiveness depending on who is given the intervention. For example, what is the cost-effectiveness of the intervention if it were delivered to only foster carers reporting high numbers of child behaviour problems compared to providing the KEEP intervention to all foster carers, regardless of the number of reported behaviour problems? Such information would be useful if there were insufficient resources to deliver the KEEP intervention to everyone. Who should be put on the waiting list? Doing this type of cost-effectiveness analysis helps to understand whether there are big or small differences in cost-effectiveness ratios depending on to whom the intervention is delivered. Doing this type of analysis requires that we have English-equivalent data. Specifically, English-specific rates of foster placement disruption for foster children aged between 5 and 12 years. However, this data was not available for the age group we needed and did not match the definition of foster placement disruption as defined in the US study.

The US study defined foster placement disruption as an index of five possible events. It includes foster placement disruption due to (i) child behaviour, either at the request of the foster carers or the caseworker deciding the child was too difficult or the caseworker deciding the child needed more intensive care (ii) the child being sent to residential care, (iii) juvenile centre, or (iv) psychiatric care and it also included (v) events of children running away. We contacted the authors to disentangle the results but the authors did not have this data.

This presented a challenge to understanding the generalisability of the US findings to the English context. To investigate, we undertook additional literature searches and met with research experts and members of the Guideline Committee (see Appendix 8.3 for detail). After consultation, it was decided that foster placement disruptions due to child behaviour and events of children running away were most generalisable to the English context. It was decided that it was unlikely that a foster placement disruption for children aged 12 years and younger would result in juvenile detention, residential care, or psychiatric care. English national data supported this (see Appendix 8.3 for detail).[29]

With this guidance from the consultation, we searched for English-equivalent data on foster placement endings as a result of child behaviour problems and rates of children running away.

[29] It is also worth noting that the US study also found that the KEEP intervention increased rates of adoption and reunification. This was also presented as an index measure, reported as a 'positive placement change', and results were not disaggregated for each outcome separately. However, after consultation with the Guideline Committee, other experts, and referring to additional academic literature, these outcomes were also considered to be less generalisable to the English context and were not included in our analysis (see Appendix 8.3 and 8.5 for more detail).

We found English national statistics regarding foster placement disruption but it did not provide the data we need. There were two sets of statistics, both of which were not robust enough for our analysis. The first reason the data were not suitable is that it counts the rate of all unplanned endings whereas we needed data on the proportion of children having an unplanned ending. Second, data were not age-specific, which is important considering that it is thought that age influences the likelihood of a placement change, and we need data for children aged 5 to 12 years. Third, the definition for 'unplanned' placement change does not match the definition we require. Data from 2014 and 2015 show a rate of 6% and 8.4% of unplanned placement changes (for children of all ages) (Ofsted 2015; Ofsted 2016). The 2014 definition of 'unplanned' placement contains three components: (i) disruptions initiated by foster carers (ii) moves to secure units, returning home to live with parents, allegations (iii) ending within 24 hours (Ofsted 2015). The 2015 definition is broken down into (i) disruptions initiated by foster carers and (ii) other reason (Ofsted 2016, p.10). If we use the more narrow definition of unplanned placement moves which is defined as the rate of unplanned endings initiated by foster carers, then the rate falls to 3.9% and 2.7% (for children of all ages), for the 2014 and 2015 years (Ofsted 2015; Ofsted 2016, p.10). Again, this definition is too narrow as it excludes caseworkers' requests for a placement ending based on child behaviour problems. Furthermore, consultation with the Guideline Committee indicated that rates of unplanned placement moves appeared to be too low. They believed that rates of planned foster placements may include disruptions due to behaviour problems.

Given that national data were not suitable, we undertook additional searches of the research literature and only found four English studies, however, data were not suitable. One was an older, non-representative cross-sectional study using case files from four local authorities. In that study, 16% of children aged 5-9 years had an unplanned placement ending in a two-year period (Farmer et al 2010).[30] In that study, no further detail is provided around the outcome of 'unplanned placement ending'. A second, older, small sample, non-representative longitudinal study used data from six local authorities based on children entering care in 1996/1997 and followed up until 2000. This study reports on the percentage of foster placements disrupting at the request of carers and at the request of children, but data are for all ages and are not age-specific. A third, older, representative, cross-sectional study used data from seven local authorities and found that foster placement disruptions were very rare among 'younger children' (as measured over a 14-month period). The study does not provide quantitative data (Sinclair et al 2003). A fourth study was a twelve-week RCT evaluating a foster carer training intervention based on a small sample with unclear representativeness. However, this study did not report on the rates of foster placement disruption (Briskman et al 2011 and Beecham et al 2012). (See Appendix 8.4 for more detail).

We also searched for English-equivalent data on the number of children running away from their placement but the data were also limited. National

[30] Sample size for the age group of 5-9 years old is not reported.

data are not age-specific. Instead, data are reported by age categories (i.e. ages 5-9 years old). Furthermore, members of the Guideline Committee and other experts believed that the national data were underestimates, even for children aged 5-9 years old. It is known that figures vary depending on the agency collecting the data, in particular, data from national statistical returns and local authority returns are underestimates and that figures from police are usually higher (see Appendix 8.3 for more detail). However, we were unable to find data based on police statistics. Furthermore, we did not find readily available data on the societal and public sector impacts of a child running away. However, a missing child is likely to incur costs to the police and social services as a result of a missing persons investigation. Additionally, we were unable to estimate the impact on QALYs as a result of a child running away.

Conclusions
In sum, our analysis estimates the resource implications of the KEEP intervention (intervention costs) relative to the short-term improvements in parenting skills and reduced child behaviour problems.

Results from the US study indicate that preventative effects are more likely to be seen among children with greater numbers of foster carer-reported child behaviour problems and for children with a greater number of previous foster placements. While we were unable to provide specific estimates of cost-effectiveness for these subgroups, it may be more cost-effective to prioritise these foster carers if there aren't enough resources to provide the KEEP intervention to everyone.

However, the entirety of the analysis and our findings need to be balanced with the knowledge that there has not been any published research of the effects of the KEEP intervention in England. Therefore, it still remains unclear whether the KEEP intervention is effective and cost-effective in the English context.

5 Linking economic evidence to recommendations

The Guideline Committee had originally recommended the KEEP intervention prior to the completion of the economic analysis on the basis of effectiveness evidence only.

Given the potentially large resource implications of implementing the intervention, NICE requested that scenario analyses be carried out, and to estimate the QALY gain and/or cost offsets required for the KEEP intervention to be cost-effective. After performing scenario analyses, the Guideline Committee believed that the KEEP intervention could be cost-effective at £20,000-£30,000 per QALY, based on the assumptions made in the scenario analyses. The assumption is that KEEP results in QALY gains of, at the very least, between 0.07 to 0.10 QALYs, further assuming that the less-resource intensive version of the KEEP intervention is provided (a cost of £2,000 per parent attending the KEEP group parenting sessions). The Guideline Committee assumed that these gains could be generated by improvements to children's levels of anxiety/depression and their ability to undertake usual

activities (key areas of the EQ-5D), based on evidence that KEEP reduced child behaviour problems and that parenting skills improved. Furthermore, the GC believed that there would be knock-on effects, as trained foster carers trained would be able to improve the quality of life of other foster care children in the household and future foster children. For further detail on the scenario analyses, please refer to NICE to request the documents.

Recommendation 1.7.12

Offer a group-based parent training intervention, for example KEEP, to foster carers of children aged 5 to 12 who have been abused or neglected and are showing problematic behaviours. Include strategies to manage behaviour and discipline positively. Provide group sessions over at least 16 weeks with groups of 8 to 10 foster carers, including video, role play and homework practice.

Recommendation 1.7.13

Consider the intervention in recommendation 1.7.12 for foster carers of children aged 5 to 12 who have been abused or neglected and are not currently showing problematic behaviours.

6 Bibliography

US RCTs

Chamberlain P, Price J, Leve LD, Laurent H, Landsverk JA, Reid JB (2008). Prevention of Behavior Problems for Children in Foster Care: Outcomes and Mediation Effects. Prevention Science, 9(1): 17–27

Chamberlain P, Price JM, Reid JB, Landsverk J, Fisher PA, Stoolmiller M (2006). Who disrupts from placement in foster and kinship care? Child abuse and neglect, 30: 409–24

Chamberlain P, Reid JB (1987). Parent observation and report of child symptoms. Behavioral Assessment, 9: 97–109

Price JM, Chamberlain P, Landsverk J, Reid JB, Leve LD, Laurent H.(2008). Effects of a foster parent training intervention on placement changes of children in foster care. Child Maltreatment, 13(1): 64–75

English national statistics

Department for Education (2014). Children's Homes Data Pack 2013. <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/388701/Childrens_Homes_data_pack_Dec_2014.pdf>

Department for Education. (2013). Data pack: improving permanence for looked after children. London: Department for Education (DfE) <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/264952/final_improving_permanence_data_pack_2013_sept.pdf>

Department for Education (DfE) (2013). Data pack: improving permanence for looked after children (PDF). London: Department for Education (DfE)

Ofsted (2016). Fostering in England, 2014-15. London: Department for Education (DfE) <https://www.gov.uk/government/statistics/fostering-in-england-1-april-2014-to-31-march-2015>

Ofsted (2015). Fostering in England, 2013-14. London: Department for Education (DfE) <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/395004/Fostering_2013-14_key_findings.pdf>

English studies on foster placement stability

Beecham J, Stevens M, Bauer A, Briskman J, Scott S (2012). Fostering Changes: Cost and Outcome Dimensions. [draft publication]

Briskman J, Castle J, Blackeby K, Bengo C, Slack K, Stebbens C, Leaver W, Scott S (2010). Randomised Controlled Trial of the Fostering Changes

Programme. Research Report DFE-RR237. Department for Education. National Academy for Parenting Research, King's College London

Farmer E (2010). What Factors Relate to Good Placement Outcomes in Kinship Care? British Journal of Social Work, 40(2): 426–44

Sinclair, I, Wilson K (2003). Matches and Mismatches: the contribution of carers and children to the success of foster placements. British Journal of Social Work, 33: 871–84

Ward H (2009) Patterns of Instability: Moves within the care system, their reasons, contexts and consequences. Child and Youth Services Review, 31 (10): 1113–18

Costs

Curtis L, Burns A (2015). Unit Costs of Health and Social Care 2015. Personal Social Services Research Unit

Meta analysis and systematic reviews

Goldman Fraser J, Lloyd SW, Murphy RA et al. (2013). Child exposure to trauma: comparative effectiveness of interventions addressing maltreatment. Rockville, MD: Agency for Healthcare Research and Quality

Oosterman M, Schuengel C, Slot NW, Bullens RAR, Doreleijers TAH (2007). Disruptions in foster care: A review and meta-analysis. Children and Youth Services Review, 29: 53–76

Rock S, Michelson D, Thomson S, Day C (2013). Understanding Foster Placement Instability for Looked After Children: A Systematic Review and Narrative Synthesis of Quantitative and Qualitative Evidence. British Journal of Social Work, 1–27

Links from placement disruption to child wellbeing

Aarons GA, James S, Monn AR, Raghavan R, Wells R, Leslie L (2010). Disentangling the Relationship between Behavior Problems and Placement Change in Child Welfare: A Multi-Level Cross-Lag Modeling Approach. Journal of the American Academy of Child and Adolescent Psychiatry, 49(1): 70–80

Newton RR, Litrownik AJ, Landsverk JA (2000). Children and youth in foster care: Disentangling the relationship between problem behaviors and number of placements. Child Abuse and Neglect, 24(10): 1363–74

Rubin DM, O'Reilly ALR, Luan X, Localio AR (2007). The Impact of placement stability on behavioral wellbeing for children in foster care. Journal of the American Academy of Child & Adolescent Psychiatry, 46(5): 610

Factors influencing adoption

Akin B (2011). Predictors of foster care exits to permanency: A competing risks analysis of reunification, guardianship, and adoption. Children and Youth Services Review, 33: 999–1011

Goerge RM (1990). The reunification process in substitute care. The Social Service Review, 64: 422−57

Leathers S, Spielfogel JE, Gleeson JP, Rolock N.(2011). Behavior problems, foster home integration, and evidence-based behavioral interventions: What predicts adoption of foster children? Children and Youth Services Review, 34: 891–9

Park JM, Ryan JP (2009). Placement and permanency outcomes for children in out of home care by prior inpatient mental health treatment. Research on Social Work Practice, 19(1): 42−51

Potter CC, Klein-Rothschild S (2002). Getting home on time: Predicting timely permanence for young children child welfare. Child Welfare, 81(2): 123−50

Selwyn J, Wijedasa D, Meakings S (2014). Beyond the Adoption Order: challenges, interventions and adoption disruption. Department for Education. Research Report. University of Bristol School for Policy Studies Hadley Centre for Adoption and Foster Care Studies

Smith BD (2003). After parental rights are terminated: Factors associated with exiting foster care. Children and Youth Services Review, 25(12): 965−85

7 Appendix

7.1 Intervention costs

In-depth description of the intervention

The intervention is delivered in groups of 3 to 10 foster carers led by a trained facilitator and co-facilitator. The intervention is manualised and comprises 16 weekly sessions. The aim is “increasing foster carers’ use of positive reinforcement relative to the amount of discipline they use” focusing on a 4 to 1 ratio and also emphasises non-harsh discipline (Chamberlain et al 2008, p.4). Manualised sessions focus on “protective and risk factors found in previous studies to be developmentally relevant and malleable targets for change” (p.4).

Group sessions lasted 90 minutes. The style of interaction is to focus on teaching concepts through group discussions, use of role-play and videotapes, rather than using teacher presentations only. Foster carers were

also given practice assignments at home relating to the weekly topic. If parents missed a meeting, they received a home visit.

The intervention was provided in a 'convenient location' and childcare was provided. The participating parents also received credit towards foster caring licensing requirements and paid parents $15 per session along with providing refreshments (Chamberlain et al 2008).

Table 8.1 – Unit costs per hour using full cost approach

Scenarios	**Facilitator**	**Co-facilitator**	**Supervisor**
Lower cost scenario: Assume facilitators are family support workers and supervisor is a social worker	£30 hour	£30 hour	£57 per hour
Higher cost scenario: Assume supervisor and facilitators are social workers	£57 per hour	£57 per hour	
Source: Curtis 2015, p.189, 194			

Table 8.2 – Costs of the direct provision of sessions

1. Direct provision of sessions, total time costs of professional input (1 facilitator, 1 co-facilitator)	**40 hours per professional**
16 weeks x 90 minutes	24 hours
Travel time, assume 1 hour per session	16 hours

Table 8.3 – Costs of training and supervision

2. Training and Supervision Total time costs of professional input: 1 facilitator, 1 co-facilitator, 1 supervisor	**109 hours per professional**
Costs of training	
5-day training, assume 8 hours per day	40 hours
Travel time, assume 1 hour per session	5 hours
Costs of supervision	
16 weeks of supervision to review and discuss videotaped sessions. Assume 90 minutes to discuss the 90 minute videotaped session. Total of 3 hours per week x 16 weeks.	48 hours
Travel time, assume 1 hour per session	16 hours

Table 8.4 – Cost of recruitment

3. Cost of recruitment[31] Total time costs of recruitment: 1 social worker		**41 hours**
Stages of recruitment	**Number of individuals**	**Time costs**
Select eligible carers from local authority database, assume 1.5 minutes to screen carers on the computer	125	3.125 hours
Contact eligible carers by telephone and letter, assume 15 minutes per person for telephone call	45/125	11.25 hours
Respond to interested carers, assume 30 minutes per person	28/45	14 hours
Number participating, assume additional 30 minute telephone call	25/28	12.5 hours

Table 8.5 – Total time costs for stages 1-3 (direct provision, training, and recruitment)

4. Professional costs for stages 1-3			
Stage	**Facilitator**	**Co-facilitator**	**Supervisor**
1. Direct provision	40 hours	40 hours	n/a
2. Training & supervision	109 hours	109 hours	109 hours
3. Recruitment	41 hours	n/a	n/a
Total hours	190 hours	149 hours	109 hours

Costs scenarios using different unit costs			
Lower cost scenario	£5,700	£4,470	£6,213
Higher cost scenario	£10,830	£8,493	
All professionals, stages 1-3			
Lower cost scenario			**£16,383**
Higher cost scenario			**£25,536**

[31] Based on recruitment described in Fostering Changes program (Briskman et al 2011, p.20).

Table 8.6 – Time costs of providing missed sessions at home

5. Costs of providing missed sessions at home Total time costs: 1 facilitator			
Total costs	**Group size 3**	**Group size 6**	**Group size 10**
Lower cost scenario Facilitator unit costs are equivalent to family support worker	£300	£375	£450
Higher cost scenario Facilitator unit costs are equivalent to social worker	£570	£712.50	£855

Calculation		
Group size of 3	**Group size of 6**	**Group size of 10**
10 hours of staff time	12.5 hours of staff time	15 hours of staff time
4 sessions missed (0.2*3 + 0.2*16)	5 sessions missed (0.2*6 + 0.2*16)	6 sessions missed (0.2*10 + 0.2*16)
6 hours direct provision 4 hours travel time	7.5 hours direct provision 5 hours travel time	9 hours direct provision 6 hours travel time
20% of total sessions were provided at home (defined as 20% of each group misses 20% of total sessions). Assume only one facilitator attends house visit. Travel time, assume 1 hour per session. Assume same duration (90 minutes) per session at home.		

Table 8.7 – Costs of other intervention components

6. Costs of other components	**Median**	**Low**	**High**
Total costs	**£2,259**	**£1,184**	**£2,718**
Childcare	£590	£202	£716
Refreshments	£519	£297	£631
Venue Hire	£995	£569	£1209
Materials	£155	£116	£162
Source: Assumed similar costs based on various group-based parenting programs (Curtis et al 2016, p.99)			

Table 8.8 – Cost of paying the foster carer for each session attended

7. Cost of incentives
£210.60 total payment to each foster carer (assuming they attended 13 sessions, based on authors reporting that 20% of sessions were missed and there are a total of 16 sessions). In the US study, the payment was $15 per 90-minute session ($10/hour), which is approximately 1.5x the hourly minimum wage. In England, national minimum wage is £7.20/hour, and we assume a payment of £16.20 per 90-minute session (£7.20*1.5*(90/60minutes)).

Table 8.9 – Total cost per foster carer, stages 1-7.

8. Total cost per foster carer			
Scenario	**Group size 3**	**Group size 6**	**Group size 10**
Lower cost scenario	£6,166	£3,140	£2,012
Higher cost scenario	£9,818	£5,038	£3,121

7.2 Making links from placement disruption to impact on child wellbeing

Three US studies that were designed to test causality found consistent results regarding the negative impact of placement instability or a high number of placement changes on subsequent child behaviour. Aarons et al (2010) provided a brief review of the literature and found that studies of association are much more frequent in the literature (using cross-sectional designs), but studies of causation are much fewer (Aarons et al 2010, p.2, 3).

Aaron et al (2010) first reviews Newton et al (2000). Their methods are based on administrative and clinical data on a moderate sized sample of n=415 individuals between ages 2-16 years old. They used data from an 18-month longitudinal foster care cohort study following children from first entry into foster care. Their findings compared those children who initially scored below the threshold on the CBCL across the three subscales to those children who scored above the threshold on at least one of the subscales. They found that (i) a higher number of placement changes is predicted by externalising behaviour problems, (ii) frequent placement changes predict negative subsequent impacts on both internalising and externalising behaviour and this was true even for those children who initially scored below the threshold.

Rubin et al (2007) uses a prospective national cohort study of N=729 children with ages between birth to 15 years. These children were followed over an 18-month period upon first entry into out-of-home care. They found that across all levels behavioural problems (as measured by the CBCL), children

experiencing instability (multiple placement moves) were more likely to have behaviour problems compared to children with early or late stability. The negative impact of unstable placements was greater for those children with initially fewer behaviour problems than those children with initially higher behaviour problems. Likewise they also found that externalising behaviour problems predicted placement moves.

Aaron et al (2010) uses the same data as Rubin et al (2007) but uses an additional 18 months of data, following up children over a 36-month period. Unlike Rubin et al (2007) the ages range from 2 to 15 years old, covering a sample size of n=422. They also wanted to identify whether age and gender modified the relationship between placement moves and child behaviour. Their findings indicated that while behaviour problems consistently predicted placement change regardless of age and gender, the findings about placement change on behaviour was less consistent. The authors conclude that the impact of placement moves on behaviour were isolated. They found that placement changes only affected subsequent externalising behaviour problems 13 months post baseline and that effect sizes were small. These were significant for the whole sample, for children ages 6-10 years old (but not for those aged 2-5 years old or aged 11+ years) and were also significant for girls but not for boys. They did not find any significant relationship of placement moves on internalising behaviour problems at 13 months nor internalising or externalising behaviour problems in the 20th and 36th months.

Aaron et al (2010) summarizes his findings within the context of previous research. Furthermore, Newton and Rubin measure the impact of placement stability whereas Aaron looks at placement moves. In relation to the limitations of their research, follow-up period may need to be longer, especially as the mean number of placement changes in the study was low (two placement changes). And this is equally true for the other two studies. Aaron et al (2010) also note that they did not stratify results according to the type of move or the number of moves that may result in different subgroup effects. For instance, whether there is a threshold effect. Likewise, whether moves are considered by the child to be positive or unwanted. The authors are also surprised that placement moves did not have a significant effect on boys' externalising behaviour problems but it did for girls. The authors try to propose some ideas regarding why there is a lack of consistent effects of placement moves on subsequent behaviour and why the effect size is small. One idea is that foster children may have actually experienced more moves with their biological family and more changes in family composition than when moving into foster care. Overall, the authors suggest that more research is needed. Relevant to our analysis, age is an important factor and children aged 6-10 years are particularly sensitive to placement moves on subsequent child externalising behaviour.

In relation to our analysis, these findings indicate that placement moves does have an impact on child behaviour but the magnitude of these effects requires additional research to understand the influence of moderating factors.

7.3 Generalisability of US findings to the English context

We asked experts to comment on the generalisability of the US findings to the English context. In particular, we asked them (1.1) do you think this would be the same in England? Should we include or exclude from analysis? (1.2) Based on available English data, do you think that these numbers should be *lower/higher* for children aged 5-12 years old?

7.3.1 Impact on foster placement disruption

<table>
<tr><th colspan="3">Residential care, juvenile centres, psychiatric care</th></tr>
<tr><td rowspan="4">National data</td><td colspan="2">These data represent all looked after children of all ages and therefore are not comparable to our sample of interest (foster care children aged 5-12 years old).
Source: Department for Education (2015), Tab A2, A3, Looked after children on March 31.</td></tr>
<tr><td>0.007%</td><td>Youth Justice Legal Status “Detained under PACE” (<5)</td></tr>
<tr><td>0.4%</td><td>Young offenders institution (270/68,840)</td></tr>
<tr><td>0.2%</td><td>NHS providing medical or nursing care (110/68,840)</td></tr>
<tr><td>GC member 1</td><td colspan="2">(1.1) Not generalisable to England.
(1.2) Unusual for children under 11. But there are very few residential homes in England in any case. Not enough beds for children nowadays. Also children placed in residential homes tend to be those who are more troubled and where placements with families have broken down</td></tr>
<tr><td>GC member 2</td><td colspan="2">(1.1) Not generalisable to England.
(1.2) It would rise between 9 and 12. Some 12 year olds would go into residential homes.</td></tr>
<tr><td>GC member 3</td><td colspan="2">(1.1) Not generalisable to England.
(1.2) No comment</td></tr>
<tr><td>GC member 4</td><td colspan="2">(1.1) Not generalisable to England.
(1.2) No comment</td></tr>
<tr><td>GC member 5</td><td colspan="2">(1.1) Not generalisable to England.
(1.2) Some going into residential homes, but numbers would be very low. But those that do go into homes are very costly.</td></tr>
<tr><td>Ian Sinclair (personal communication)</td><td colspan="2">“In England, for example, it would be highly unusual for a child of under 11 to be placed in a psychiatric establishment or juvenile correctional facility… as is the use of residential children's homes (which do have very high costs).”</td></tr>
</table>

Foster placement changes	
National statistics & English research	See Appendix 8.4 for more detail.
GC member 1	(1.1) Very likely that results are generalisable to England. (1.2) Underestimates
GC member 2	(1.1) Very likely that results are generalisable to England. (1.2) No comment
GC member 3	
GC member 4	
GC member 5	(1.1) Very likely that results are generalisable to England. (1.2) Underestimates. Research studies will have higher estimates because of the level of engagement. This discrepancy is widely recognized.
Ian Sinclair (Personal communication)	"In England… Placement breakdowns are also relatively rare among those under 11..."

Child absconding	
National data	0.36% probability that a child goes missing from foster care aged 5-9 years old, in a one-year period. Source: Department for Education (2015, Tab G1) <u>Frequency that a child goes missing from foster care (all ages)</u> 1 65% 2 17% 3 7% 4+ 10% Source: Department for Education (2014). <u>Percent distribution of durations that children go missing (all ages)</u> <24 hours 50% 1-6 days 36% 7-28 days 11% 28+ days 3% Source: Ofsted 2016, p.16 <u>Distribution of reported reason for going missing, number of times (all ages)</u> Contact with family or friends, 8,480 (50%) At risk of child sexual exploitation, 1,250 (7%) All other known reasons, 3,115 (18%) Reason unknown, 4,240 (25%) Total, 17,085 (100%)

	Source: Ofsted (2016)
Child absconding	
GC member 1	(1.1) It is very likely that results are generalisable to England.
GC member 2	(1.2) Underestimates
GC member 3	(1.1) It is very likely that results are generalisable to England.
GC member 4	(1.2) No comment
GC member 5	

7.3.2 Impact on reunification or adoption

Reunification	
National statistics	Data are seriously limited for our purposes. We do not have age-specific data. We only have data for all looked after children of all ages. Source: Department for Education (2015), Tab A1, D1. Total reunified = 10,300 Total ceasing = 30,430 Total looked after = 68, 840 Probability of being looked after = 15% (10, 300 / 68,840)
GC member 1	(1.1) Generalisability to England is unclear. Most likely not generalisable.
GC member 2	
GC member 3	
GC member 4	(1.1) Generalisability to England is unclear. Most likely not generalisable. However, difficult to say the extent to which behaviour drives those rates, if at all. Based on experience in the court, improvements in child behaviour may be a reason to not return the child to home because it tends to indicate how bad things were at home and which impacted on the child's behaviour. Indeed, improvement in behaviour in foster care may well actually reduce reunification.
GC member 5	(1.1) Generalisability to England is unclear. Most likely not generalisable. Reunification can also be defined as a permanent or, in the case of a care order; they can be placed with their parents under specific regulations. Return to birth parents under care orders often breaks down.

Adoption	
National statistics	Data are limited for our purposes because they are not completely comparable for our sample of foster care children aged 5-12 years old. Adoption: Adopted, Looked after children, ages 5-9 years = 960 All looked after children aged 5-9 years = 13,920 Probability of adoption = 6.89% Special guardianship: Total number with special guardianship = 3,330; total number looked after = 68,840. All aged children, probability of special guardianship = 4.8% (3,330/68,840). Source: Department for Education (2015), Tab A1, E1.
GC member 1	(1.1) Look at other resources first, and then decide. (1.2) No comment.
GC member 2	(1.1) Generalisability to England is unclear. No association with adoption. (1.2) No comment.
GC member 3	(1.1) Look at other resources first, and then decide. Does not anticipate a relatively strong influence on adoption. However, there is an association. Agrees with Harriet that increased placement moves may increase risk of not being adopted but would not like to say whether this is anything more than a small percentage. The approach to selecting appropriate cases for adoption is in a state of flux with tension between the approaches of the court and of the government (i.e. there has been a sharp reduction in adoption orders and a corresponding increase in orders for reunification or for placement under special guardianship orders with extended family placement from courts. The reason is not clear; there may be a different approach in the court or perhaps social workers are misinterpreting the view of courts and so they may not be planning for adoptions). The Court may be tending to a more European approach, emphasising that adoption is the last resort needing recognition of the right to family life and the need to justify state intervention, so that there is more emphasis on reunification. This may affect social work practice and reduce the planning for adoption. Traditionally the social worker has focused on the best solution in the interests of the child in social work terms. In any case, there is this state of flux as to the relationships between child behaviour and adoption and reunification. Additionally it should be noted that at the same time, some foster placements become adoptive placements so as perhaps to blur a distinction between the two.
GC member 4	(1.1) Look at other resources first, and then decide. There are many other stronger factors associated with adoption. However, agree that increased placements indicate more serious behaviour problems and that this may decrease likelihood of placement.
GC member 5	(1.1) Look at other resources first, and then decide. Children with many placements would likely to have behaviour problems and therefore less likely to be adopted. More likely to find data looking

	at adopted children (and the number of placements or levels of child behaviour). However, not that many 5-12 year olds go into adoption. Perhaps it is best not to include SGOs and adoption – we don't know if it's related to parenting. At the same time there may be a link through frequency of placement changes and impact on adoption.

7.4 Rates of foster placement disruption in England

The aim of this appendix is to identify English baseline data on foster placement disruptions. We searched for comparable English data and assessed whether it is sufficiently comparable to the US study's sample, definition of placement disruption, and whether English data are representative and recent.

Source	**Does the definition match the US RCT**	**Is the sample comparable to the US RCT (age)**	**Is the data representative and recent**	**Time horizon**	**Value**
US RCT Chamberlain 2006, 2008, Price 2008	**This is the original study**		**1999-2004**	**6.5 months**	Intervention 14.3% Control 12.2%
English national Statistics Department for Education (2015) Cross-section	**Over-estimate** "Total placement changes" includes planned & unplanned moves, positive & negative reasons. Does not include children absconding.	**Not an exact match.** Age groups include those aged 5-9 and 10-15 years old.	**2011, national statistics**	**12 months** 2 placements 3 placements 4/5 placements 6-9 placements 1 placement 1+ placements	 17-22% 5-6% 1-4% 0-1% 70% 30%
				We combined results for age groups 5-9 and 10-15 years old as figures were similar.	
English national Statistics Ofsted 2016 (p.10, 15) Cross-section	**Underestimate** (2016 definition different to 2015) "Unplanned endings" endings not included in social work plan either in ending or timing	**Not an exact match.** All ages (0-16)	**2014/15, national statistics**	**12 months** Total unplanned endings, n=7,245 of N=85,890 kids in foster care. Of which, 47% initiated by foster carers (n=3390)	 8.4% rate of unplanned placement endings 3.9% 4.5%

	(i) initiated by foster carers (ii) other reason			Of which, 53% occurred for another reason (n=3855) Of all unplanned endings, 20% occurred within 24 hours.	
English national Statistics Ofsted 2015 (p.9) Cross-section	**Underestimate** (2015 definition has greater number of reporting categories) "Unplanned endings" endings not included in social work plan either in ending or timing (i) initiated by foster carers (ii) ending within 24 hours and (iii) children/young people requiring secure placements; returning home to live with parents; and allegation	**Not an exact match.** All ages (0-16)	**2013/14, national statistics**	**12 months** Total unplanned endings, n=5,240 of N=84,450 kids in foster care. Of which, 44% initiated by foster carers Of which, 34% due to secure placements; returning home to live with parents; and allegations Of which, 22% ending within 24 hours.	6% rate of unplanned placement endings 2.7% 2.1% 1.2%
English study Farmer et al (2010) Cross-section	**Unclear perhaps similar definition** "Unplanned endings" (no further detail)	**Not exact match.** All ages (0-16)	**Older, non-representative data** - 4 local authorities, n=270 case files selected (unclear selection process); unclear date	**24 months** Age at the time of study placement 0-4 years **5-9 years** 10+	 12% **16%** 37%

English study Briskman et al (2010) Beecham et al (2012) RCT "fostering changes"	**Unclear**	**Very close match.** Range 2-12 years Mean 7.9 years, SD=3.1	**Recent (2009), small sample, unclear representativeness.** N=63 carers, 89 foster kids, 4 local authorities	**12 weeks** Service use patterns and rates of unplanned endings were not reported although social care service use data were collected and costs were reported.	
English study Sinclair (2003) Cross-section	**Unclear.** Definition of breakdown "when foster carer, family placement or child social worker said that this [breakdown] had happened."	**Not an exact match.** All ages (4-16)	**Older, representative data.** 7 local authorities, nationally representative but under-representing kinship placements.	**14 months** "Moreover disruption was very rare among younger children and so difficult to predict" (p.877)	
English study Ward et al (2009) Longitudinal study	**Over-estimate.** All placement moves.	**Depends.** Some info provided for similar match (ages 5-9 years) others for whole sample (all ages).	**Older, small, unrepresentative study.** N=242, 6 local authorities, came into care between April 1996/97 until 2000. Eligibility: looked after for min. 12 months.	<u>Aged 5-9 years old, 3.5 year period</u> 1 placement 2 placements 3-5 placements 6-9 10+ <u>All ages, 3.5 year period, number of moves and reason (n=700 moves)</u> <u>Foster carer</u> Carer-initiated disruption Foster carer required relief Return from relief placement <u>Child-related</u> Child requested disruption Child absconded	 19% (n=12) 23% (n=15) 40% (n=26) 19% (n=12) 0% 21% 7% 7% 5% 3% 3%

				Return from absconding <u>Moves initiated by local authority</u> Planned transition Placement no longer available ---------- Moved out of looked after care (n=125 moves out of care /242 long-stay children)	 43% 11% ---------- 52%

Ward et al (2009)				
Placement number	**Number of placements**	**% Ending, planned**	**% Ending, disruption or absconding**	**No ending**
1	242	57%	15%	28%
2	197	37%	19%	44%
3	142	31%	20%	49%
4	97	30%	24%	46%
5	70	33%	19%	46%
6	53	34%	19%	47%

7.5 Factors influencing adoption, two US studies

In our search of the literature, we did not find English-specific studies regarding the factors that influence adoption. We did find two recent US studies (Akin et al 2011 and Leathers et al 2011).

Regarding the links from number of placements and likelihood of adoption, Akin et al (2011, p.1001) conducted a brief review of the literature and found that placement stability was rarely used as a predictor when measuring impact on permanent exits. The findings were conflicting. Two studies found no relationship (Akin et al 2011 citing Park & Ryan, 2009; Potter & Klein-Rothschild, 2002) and two other studies found a significant association between lower rates of reunification and adoption and increased number of placement moves (Akin et al 2011 citing Goerge 1990 and Smith 2003).

Akin et al (2011) then conducted their own analysis and provided hazard ratios for the association between certain variables and likelihood of adoption. The strength of their analysis is that it is based on a large cohort study (N=3,351) followed up between 30-42 months. The authors state that the limitations of their research are that, first, it is not meant to determine causality but rather associations; second, they could not examine re-entry into care; third, they did not take other variables into account that might be influential,

for example role of caseworkers, agencies, communities, and courts; and fourth, more research is needed to corroborate findings.

Akin et al (2011) found that early placement stability increases the likelihood of adoption but their definition is different than the one needed for our purposes. Their definition is narrower, which defines early placement stability as 0-2 placements within 100 days versus having 3 and greater placements by the 100th day. This is useful information, but their study looks at those entering care whereas our US study looks at a cross-section of children who have had varying lengths of stay in care. Therefore, the results are not immediately transferrable.

They also identified that having mental health problems reduces the likelihood of adoption. However, it is not comparable to our definition as we look at child behaviour problems. Their definition is based on a categorical variable defined as having or not having socio-emotional difficulties.

They also found that factors reducing the likelihood of adoption include running away events and increasing age. However, having a physical disability increases the likelihood of adoption relative to mental health difficulties. Non-kinship foster care and intact sibling placements were also found to increase the likelihood of adoption.

The second and final study that we identified, Leathers et al (2012) reused data from an adaptation of the KEEP intervention study (Price et al 2008) but the study is limited due to its very small sample size (N=31). Given the small sample size, confidence in the findings is severely limited. The authors analysed the whole sample and found that “externalizing behaviour problems had a negative effect on both foster home integration and adoption, and foster home integration had an independent positive effect on adoption. Internalizing behaviour problems (e.g., depression/anxiety) were not related to adoption or integration.” While this finding is helpful, it is again unclear whether findings are applicable to the English context. For example, in the US, foster-care placements that become adoptions occur much more frequently (56%) than they do in England (15%) (Selwyn et al 2014, p.17)

In summary, it is not clear that US findings about adoption rates are generalisable to the English context.

Child abuse and neglect

Economic Appendix C3.4

Cost-effectiveness analysis

SafeCare intervention – Home visiting intervention for maltreating parents and their biological children aged between 2 and 12 years old

Review question 15

What is the impact of social and psychological interventions responding to abuse and neglect?

This report was produced by the Personal Social Services Research Unit at the London School of Economics and Political Science. PSSRU (LSE) is an independent research unit and is contracted as a partner of the NICE Collaborating Centre for Social Care (NCCSC) to carry out the economic reviews of evidence and analyses.

Authors of the report:

Marija Trachtenberg and Jennifer Beecham

Contents

1 Background and aims

The aim of this report is to undertake a cost-effectiveness analysis on the SafeCare intervention described in recommendation 1.6.8.[32] This intervention was selected for further economic analysis due to the lack of cost-effectiveness information suitable to inform UK policy and practice and it was agreed with the Guideline Committee.

The aim of the SafeCare intervention is to reduce the recurrence of maltreatment, as measured by re-reports to child protective services (substantiated and unsubstantiated reports).

SafeCare is a home visiting intervention provided on a weekly basis for 6 months with sessions lasting between 60 to 90 minutes. Two types of SafeCare were trialled, coached and un-coached versions. In the coached version of SafeCare, the coach accompanies the home visitor on a monthly basis and provides help with logistics (it is not meant to improve home visitor fidelity to the SafeCare model). The US study does not describe who provides home visiting or who provides coaching but states that they have minimum workforce qualifications. We assume that, with appropriate training, a range of professionals might be able to provide the SafeCare intervention. A potential home visitor in the UK could be a family support worker, health visitor, or children's social worker.

The intervention is compared to two active comparison groups: coached and uncoached home visiting.

Similarities between the intervention and comparison groups are that both intervention and comparison groups, home visitors have caseloads between 17 and 18 families. Home visits last between 60 and 90 minutes and provided for 6 months. Both provide weekly visits, involve service goals, use case management practices, have similar reporting requirements, and use of similar assessment tools. Both services are funded in the same manner (Chaffin et al 2012, p.510).

The difference between the services is the SafeCare approach to the home visit. It is a manualised, structured behavioural skills training model that address "parent-child or parent-infant interaction, basic caregiving structure and parenting routines, home safety, and child health" (Chaffin et al 2012, p.511).

The sample characteristics included nonsexual abusive/maltreating parents of children aged up to 12 years who have been referred by child protective services (Chaffin et al 2012, p.510).

[32]This intervention was founded on evidence identified in review question 15. Review question 15 looks at the impact of interventions responding to children and young people who have experienced abuse and neglect.

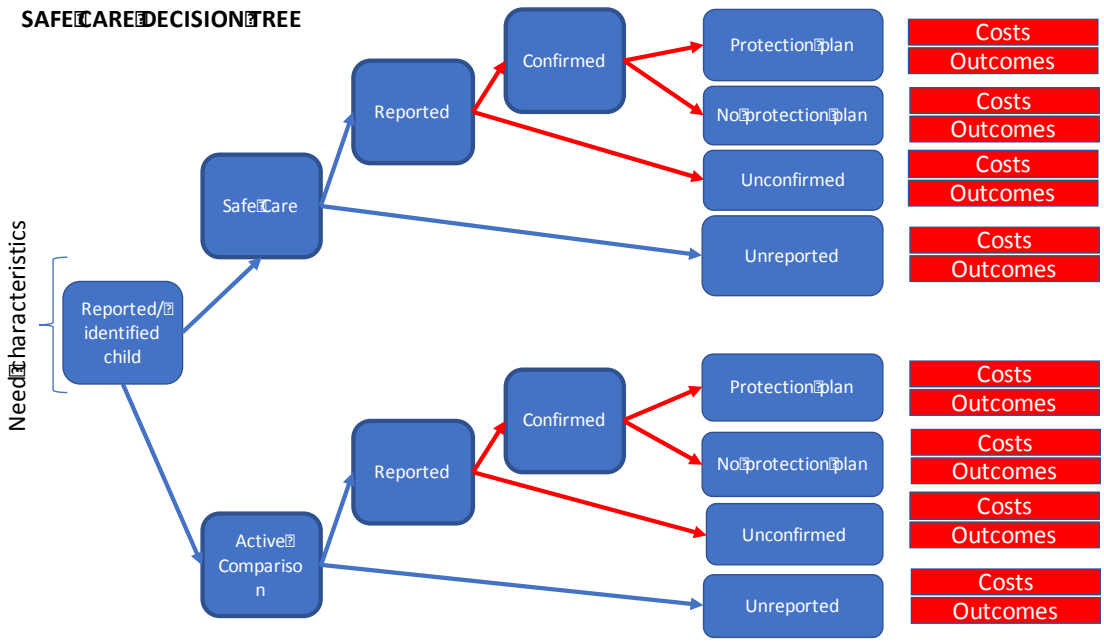
SafeCare was evaluated in a single, large sample, US RCT (n=2,175) between 2003 and 2006 (Chaffin et al 2012). Treatment completion rates were high, 89% and 87% for intervention and comparison groups respectively. This study was identified in a good quality systematic review (Goldman Fraser et al 2013). It rated the US study as being of moderate quality. The rationale for the rating is not provided.

It is important to note that, as this is a US study, we must be cautious about generalising the findings to the English context. If standard social care services in the US are poorer than compared to standard care in England, then the impact in England may not be as large. Such findings have occurred elsewhere when replicating US social care interventions in England. While we cannot know for sure until research is conducted in England, this issue is worth noting.

2 Methods

The initial aim of this report was to model the reported outcomes of the SafeCare intervention (re-reports to child protective services) to impacts on QALYs and public sector costs. However, we were unable to do so because this requires us to make assumptions about what happens to the child after the report. The results from the US study do not distinguish between reports that are confirmed or unconfirmed. Therefore, any modelling using this outcome would introduce too much uncertainty. This is illustrated in Figure 1, where the red arrows and red boxes indicate that a majority of the parameters and causal pathways that lacked evidence required for developing a decision model. Trying to fill in these gaps would require assumptions which would be difficult to substantiate, and therefore would render the analysis unhelpful for decision makers.

Figure 1 – Structure of a decision model to evaluate the impact of the SafeCare intervention on QALYs and public sector costs



Therefore, our cost-effectiveness analysis only calculates the additional resource implications of the SafeCare intervention relative for the improvements in the reported outcome. The time horizon of the analysis is the same as in the US study: intervention costs accrue over a 6-month period but the outcomes are measured over a 6-year time horizon.

The reported outcome, the relative risk reduction in the rates of re-reports to child protective services, is based on US data, which is unlikely to be similar to baseline data from the UK.[33] The relative risk (RR) is the ratio of the probability of an outcome occurring for the intervention group relative to the probability of it occurring in the comparison group (see equation 1).

Equation 1:

$$RR = \frac{Intervention: \ probability\ of\ re-report\ to\ child\ protective\ services}{Comparison: probability\ of\ re-report\ to\ child\ protective\ services}$$

The information on RR is more useful when we can approximate UK-specific impacts. To do this, we need UK baseline data on the rates of re-report to child protective services specifically for maltreating parents of children aged 2 to 12 years. However, we were unable to find suitable English-equivalent data for our purposes. Had we been able, we might be able to approximate the reduction in the number of children with a re-report to child protective services, which is more useful to decision makers than information on relative risk alone.

To do this, we searched for data using national statistics from the Children in Need returns. However, data from the Children in Need returns only provide statistics on referral rates for children of all ages (they were not disaggregated) and statistics do not distinguish between children with at least one previous referral and those with their first-ever referral. The nearest approximate statistic is the proportion of all referrals which are re-referrals – however, this data is not suitable for our purposes as it is for children *of all ages* and it only counts those who have been referred within 12 months of a previous referral (Department for Education 2016, p.5).[34] This statistic is not suitable because it excludes children who have been referred more than 12 months ago. Given the lack of data, our cost-effectiveness results are

[33] Munro and colleagues (2011) compared rates of referrals to child protective services between the USA and England and found that rates of referrals in the USA may be lower than that in England. In contrast to the USA, referrals to child protective services in the England are for both requesting services for children in need and for those based on suspected child maltreatment. In the USA, referrals usually do not include those requesting services for children with lower levels of need as they are usually referred onto the voluntary sector (Munro et al 2011, p.34). As of 2010, referral rates in the USA were 44.1 per 1,000 children compared to England with 53.9 per 1,000 children (Munro et al 2011, p.35).

[34] In 2015, there were 635,600 referrals for all children aged 0 to 17 years and this is similar to previous years. This is a rate of 5.5 children per 100 (Department for Education 2015, p.5). Of these, 24% were re-referrals, defined as those re-referred within 12 months of a previous referral, this is a rate of 1.3 per 100 children (Department for Education 2015, p.6).

reported using the outcome of a relative risk reduction in re-reports to child protective services (rather than the actual numbers of reductions).

3 Results

3.1 Parameters used in the analysis

Effects
The authors of the US study analysed intervention effectiveness using survival analysis using two different statistical approaches. It is worth noting that the study only provides pooled effectiveness results for the two types of SafeCare versus the two types of the active comparator (coached and un-coached home visiting) (Table 1). Said another way, they did not provide separate results for coached and un-coached versions of SafeCare versus coached and un-coached comparison services.

Using the first statistical approach (the 4-strata solution), SafeCare results in statistically significant reductions in any report to child protective services (hazard ratio = 0.83, 95% CI=0.70–0.98). The second statistical approach (using coarsened exact matching) has similar results but a smaller confidence interval (hazard ratio, 0.83, 95% CI=0.75 – 0.93). In our analysis, we use the results with the larger confidence interval to keep our estimates of cost-effectiveness conservative.

Most importantly, the effects of SafeCare are sustained beyond the 6-month period of the intervention. There was a *sustained* reduction in reports to child protective services lasting the next five-and-a-half years.

Table 1 – Effectiveness of SafeCare on reducing the risk of report to child protective services (over a 6 year period).

Pooled effect	Statistical approach	
	4-strata solution	Coarsened exact matching
SafeCare (coached & un-coached) vs. Comparison (coached & un-coached)	Statistically significant effect HR = 0.83 CI = 0.70–0.98 P = 0.03	Statistically significant effect HR = 0.83 CI = 0.75–0.93 P = 0.001

Costs
Our cost analyses are based on national average unit costs and use a full-cost approach, which is in line with accepted practice (Curtis 2016). All costs reflect the 2015/16-price year.

Our analysis only includes intervention costs. As the study only presents the pooled effects for coached and un-coached interventions, this presents some challenges for the Guideline Committee regarding whether to recommend the

coached and un-coached interventions. For our cost-effectiveness analysis, we use the average cost of the coached and un-coached intervention. Respectively, the range of intervention cost estimates is between £3,500 and £6,000 per family (delivered over a 6-month period). The cost of the intervention is influenced by several factors.

First, the study does not describe who provides home visiting but states that they have minimum workforce qualifications. We assumed that a potential home visitor in England could be a family support worker, health visitor or children's social worker. Respectively, the hourly unit costs for each are £45.00, £76.00, and £85.50; this includes an assumption that an additional 30 minutes of administrative family-related work was also required. We chose those professionals to capture the lower and upper cost estimates. Second, we assumed that coaching and training would be delivered by a child social worker – the rationale was to provide conservative estimates of intervention cost. However, in practice, anyone who is qualified to deliver training is eligible, and could also be a family support worker or health visitor. These 3 assumptions about who delivers the intervention are the main factors that influence the range of intervention cost estimates. These estimates are presented in Table 2.

We also make additional assumptions. The authors of the US study report that home visits occur weekly and last between 60 to 90 minutes, our cost estimates assume an average visit of 75 minutes.

Our cost estimates also include the costs of travel, which we assumed to be 1 hour for each home visit.

We also include the costs of training. Group training is provided to 3 to 4 home visitors, for 5 days, and we have assumed a day's training lasts 8 hours per day. Our analysis conservatively assumes a group size of 3. The home visitor also receives nine directly observed field sessions as a part of training, and we assume this lasts an average of 75 minutes.

Details on the calculation of intervention costs are provided in Appendix 8.

Table 2 – Total cost per family

Intervention cost scenarios	Lower	Middle	Upper
Cost per family used in the analysis (Average of coached & not coached)	£3,500	£5,400	£6,000
Not coached	£3,000	£4,900	£5,400
Coached	£4,100	£6,000	£6,600
Note: Costs have been rounded to the nearest 100			

3.2 How the results were calculated

Our results are calculated and presented in 2 different but complementary ways, both of which help in deciding whether the KEEP intervention is cost-effective.

In the first approach, we present the findings as a mean cost-effectiveness ratio (Table 3), defined as the mean cost for a 1-unit reduction in an outcome. In this analysis, we present the outcome of a 1% reduction in a report to child protective services (Table 3).

The cost-effectiveness ratio is useful when you must select between different interventions that measure the same outcome. A decision maker would want to choose the intervention that has the lowest cost per unit of improvement. A decision-maker may also have a limit as to how much they are willing to spend for an additional unit of effect. This is termed the 'cost-effectiveness threshold'.

The only outcome for which there is an established threshold is the quality-adjusted life-year (QALY). The QALY is a measure of health-related quality of life. NICE guidance suggests that interventions with cost-effectiveness ratios equal to or less than £20,000 to £30,000 per QALY are cost-effective, in areas where the QALY is applicable. However, for non-QALY outcomes, there is no guidance around an acceptable cost-effectiveness ratio. This means there are no established thresholds for determining the cost-effectiveness of an intervention based on the outcomes of a 1-unit reduction in a report to child protective services, unless there was some way of linking this to QALYs. However, as we mentioned earlier in the section on methods, we were unable to find evidence that makes these links. Had we been able, we could compare the SafeCare intervention to any other intervention and any other population so long as results are measured with QALYs.

In the second approach, we present the results using a cost-effectiveness acceptability curve (CEAC) (Figure 1 and Figure 2). The CEAC summarises the parameter uncertainty surrounding the intervention's cost-effectiveness ratio. Parameter uncertainty, in this case, relates to the variation in effectiveness, which considers the distribution of effect, rather than using only information about the mean effect. With that information, it is possible to provide a probabilistic estimate of the intervention's cost-effectiveness. We undertook a probabilistic analysis using a simulation, in particular, a Monte Carlo simulation. In our analysis, we only simulated outcomes, and not costs. We did not use probabilistic sensitivity analysis for costs because intervention inputs are fixed, and not probabilistic. As there is a range factors influencing intervention costs, it is more appropriate to use scenario analyses, as we have done and have presented earlier. In our analysis, we only simulated outcomes, and for three intervention cost scenarios using the lower, middle, and upper estimates described earlier (£3,500 and £5,400 and £6000 per family, rounded to nearest hundred).

The probability of the intervention's cost-effectiveness is calculated in several steps. First, we create a hypothetically large sample size (we chose 1,000, as this is standard practice). We then created a randomisation formula that picks a value, at random, within the constraints of the mean and standard deviation for the outcome. In the absence of information about the distribution of reports to child protective services, we assumed that the chances of getting different numbers are based on a 'normal' distribution. The 'normal' distribution assumes that chances of getting different numbers on either side of the mean is symmetrical, and that numbers closer to the mean have a higher chance of occurring than values farther away from the mean. Effectively, we are assuming that a majority of the sample will have an effectiveness outcome that is close to the mean score, and a smaller percentage of the sample will have effects that are either smaller or larger than the mean.

When the probabilistic analysis is set up, it is able to tell us the probability that the intervention is cost-effective at a certain amount of money. For example, a decision maker may only be willing to pay £20,000 per QALY. The CEAC tells you the probability that the intervention is cost-effective at various amounts that a decision-maker is willing to pay. This is calculated by finding the percentage of times that the intervention is cost-effective at or below £20,000, and this is done over a large number of simulations (in our case, 1,000 simulations). For example, we might find that 75% of the 1,000 simulations result in a cost-effectiveness ratio equal to or below £20,000. However, as discussed earlier, when we have cost-effectiveness ratios with non-QALY outcomes, there is no established threshold for assessing cost-effectiveness. In any case, presenting results using a CEAC is still helpful because it summarises the uncertainty surrounding the intervention's cost-effectiveness ratio.

3.3 Results: mean probabilistic cost-effectiveness ratio

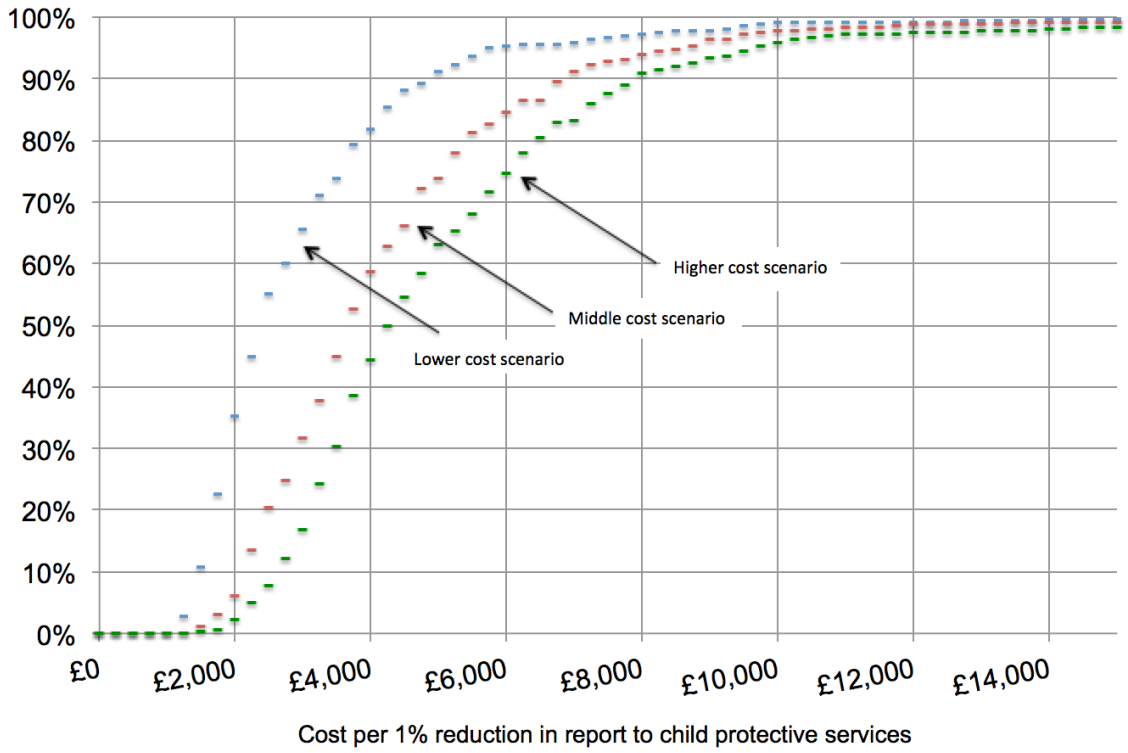
The mean effect of the SafeCare intervention is to reduce the rate of re-reports to child protective services by 17% (95% CI=0.70–0.98). When this information is simulated probabilistically alongside our intervention cost estimates, this results in a mean probabilistic cost-effectiveness ratio between £286 (in the lower intervention cost scenario) and £490 per 1% reduction in reports to child protective services (in the upper-intervention cost estimate) (Table 3).

Table 3 – Mean probabilistic cost-effectiveness ratios

Intervention cost scenarios	Lower	Middle	Upper
Mean probabilistic cost-effectiveness ratio for a 1% reduction in a report to child protective services	£286	£430	£490

In relation to the CEAC, the simulated data show that there is a 95% probability that the additional cost of the SafeCare intervention, compared to the active comparison group in the US study, is less than £5,750 for a 1% reduction in reports to child protective services (in the lower-cost intervention scenario). In the middle-cost scenario, this corresponds to £8,250, and in the higher-cost scenario, this corresponds to £9,500 (Figure 1).

Figure 1 – Cost-effectiveness acceptability curve: probability of cost-effectiveness for a 1% reduction in reports to child protective services



4 Discussion

Findings

This analysis estimated the English-equivalent costs of the intervention per family, assuming the average cost of the coached and not coached versions. Three different cost scenarios were presented to reflect the different types of professionals who could potentially deliver the intervention (family support worker, health visitor, child social worker). Across each of the scenarios the cost per family for 6 months of home visiting is between £3,500 and £6,000. The US study found that the intervention led to an average 17% relative risk reduction (95% CI, 2% to 30%) in any report to child protective services, which was sustained over the next 5.5 years.

It is very important to note that this does not imply that the lower-cost version of the intervention (uncoached) is more cost-effective than the higher cost version (coached). The study was not designed to answer this question. We

are only able to estimate the potential range of intervention costs, but we do not know whether this has any effect on outcomes.

Strengths and limitations

Our analysis is useful but limited for several reasons. First, we could not compare the intervention's relative cost-effectiveness to other interventions with similar objectives. One such example is parent-child interaction therapy, which forms the basis of Recommendation 1.6.9. While this intervention had similar objectives we could not include it in our analysis because of differences in sample characteristics. In one study the sample included only physically abusive parents with children present in the home (Chaffin et al 2004). The second study included parents with histories of abuse or neglect and 66% of families had children removed from the home but parents still had legal parental rights (Chaffin et al 2011). Even if samples were similar, it would still have not been possible to do a comparison. This is because the studies do not share a common comparison group, which is a pre-requisite for combining results across different studies (using indirect treatment analysis).

The second limitation is that we could not model the additional impacts of a re-report to child protective services on QALYs or on societal and public sector costs due to the lack of evidence about these key causal links. Therefore, we must emphasize that without a robust economic evaluation in the UK, we cannot be sure whether or not the intervention is cost-effective, based on currently available evidence.

The third limitation is that results are based on US research. This is important because US and English 'standard care' services are likely to be different. In this study, standard care was 'active', providing home visiting services of equal length and duration, with exception of the SafeCare approach to the home visits. If English standard care services are better than standard care services in the US, then the effects in England might be smaller.

Recognising these limitations, this is still a first step in estimating the cost-effectiveness of the SafeCare intervention using English-equivalent intervention costs. Research on the SafeCare model is needed in England. This research should include an economic evaluation, which includes parents' and children's use of services across health, social care, education, and criminal justice sectors. Research should also include impacts on individuals' outcomes, such as parents' and children's physical, health and overall wellbeing. Research should also be long enough so that lagged effects are captured and to understand whether effects are sustained beyond the intervention period.

5 Linking economic evidence to recommendations

The Guideline Committee had already recommended the SafeCare intervention prior to the economic analysis. The Guideline Committee still decided it was worth recommending the SafeCare intervention based on the results of the study's reported effectiveness. Given the potentially large resource implications of implementing the intervention, NICE requested that

scenario analyses be carried out, and to estimate the QALY gain and/or cost offsets required for the SafeCare intervention to be cost-effective. After performing scenario analyses, the Guideline Committee believed that the SafeCare intervention could be cost-effective at £20,000-£30,000 per QALY, based on the assumptions made in the scenario analyses. The assumptions are that SafeCare is able to generate cost-offsets of £35,000 and an average of 1.44 QALY gains over the child's lifetime (at a threshold of £20-30,000 per QALY). This is assuming the provision of either the lower- or higher- resource intensive provision of SafeCare (between £3,500 and £6,000 per family). This is also based on the Guideline Committee assumption that 50% of reports to child protective services are confirmed cases of abuse and that social services responds to 50% of these cases, of which it is assumed that they are able to stop abuse within a year. However, the Guideline Committee believed that SafeCare would be more cost-effective than described, as they believed that the cost-offsets were too low. For further detail on the scenario analyses, please refer to NICE to request the documents.

Apart from this, we pointed out to the Guideline Committee that the results of the SafeCare intervention are based on a pooled effect of coached and un-coached SafeCare home visiting. The purpose of coaching is to help home visitors resolve logistical issues and was not used as a tool to increase fidelity to the SafeCare intervention. We did not have information as to whether effects were different, depending on whether coached or un-coached versions of SafeCare were provided.

This presents some challenges around making a recommendation on SafeCare and the interpretation of the cost-effectiveness analysis. Given that we presented the average cost of the SafeCare intervention based on the costs of both coached and un-coached versions, in the worst-case scenario, we may have slightly over-estimated the intervention's cost-effectiveness. The additional cost of coaching is £1,100 per person, meaning that, if the coached intervention were provided, we have over-estimated cost-effectiveness by £550.

However, this may be a small issue if we consider that our entire analysis may be underestimating the intervention's benefits. We were unable to quantitatively capture the potentially positive consequences of preventing a re-report to child protective services. For example, the identification of children with substantiated report of abuse and neglect reduces the duration that the child experiences abuse and which leads them to receive appropriate care and services, which we hope improve their outcomes.

Recommendation 1.7.10

Consider a comprehensive parenting intervention, for example SafeCare, for parents and children under 12 if the parent or carer has physically or emotionally abused or neglected the child. This should be delivered by a professional trained in the intervention and comprise weekly home visits for at least 6 months that address:

- parent–child interactions
- caregiving structures and parenting routines
- parental stress
- home safety
- any other issues that caused the family to come to the attention of services.

As part of the intervention, help the family to access other services they might find useful.

6 Bibliography

Chaffin M, Hecht D, Bard D, Silovsky JF, Beasly WH (2012). A statewide trial of the SafeCare home-based services model with parents in child protective services. Pediatrics, 129 (3): 509–15

Chaffin M, Funderburk B, Bard D, Valle LA, Gurwitch R (2011). A Combined Motivation and Parent–Child Interaction Therapy Package Reduces Child Welfare Recidivism in a Randomized Dismantling Field Trial. Journal of Consulting and Clinical Psychology, 79(1): 84–95

Chaffin M, Silovsky JF, Funderburk B, Valle LA, Brestan EV, Balachova T, Jackson S, Lensgraf J, Bonner BL (2004). Parent–Child Interaction Therapy With Physically Abusive Parents: Efficacy for Reducing Future Abuse Reports. Journal of Consulting and Clinical Psychology, 72(3): 500–10

Curtis, L, Burns A (2015). Unit Costs of Health and Social Care 2015. Personal Social Services Research Unit

Department for Education. (2015). Characteristics of children in need: 2014 to 2015. SFR 41/2015 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/469737/SFR41-2015_Text.pdf

Goldman Fraser J, Lloyd SW, Murphy RA et al. (2013). Child exposure to trauma: comparative effectiveness of interventions addressing maltreatment. Rockville, MD: Agency for Healthcare Research and Quality

Munro ER, Brown R, Sempik J, Ward H, Owen C (2011). Scoping review to draw together data on child injury and safeguarding and to compare the position of England with that in other countries. Research Report DFE-RR083. Department for Education

7 Appendix: costing the intervention

Table 8.1 – Unit costs per hour using full cost approach

Scenario	Home visitor	Supervisor/model developer
Lower cost scenario: Assume home visitors are family support workers and supervisor is a social worker	£45 per hour	£76 per hour
Middle cost scenario: Assume home visitors are 'health visitors' and supervisor is child social worker	£76 per hour	
High cost scenario: Assume home visitor and supervisor are child social workers	£85.5 per hour	
Note: Unit costs include 30 minutes of patient-related work.		
Source: Curtis and Burns 2015, pp171, 189, 194		

Table 8.2 – Time costs for home visiting sessions, per service user

Costs of home visiting sessions, per service user	Supervisor (coaching)	Home visitor
	15.75 hours	58.5 hours
6 months assuming a 75-minute visit. Home visitor, weekly basis (26 weeks) 52 weeks ÷ 2=26 weeks Supervisor coaches, monthly (7 weeks) 26 weeks ÷ 4=6.5 weeks, rounding up to 7	8.75 hours	32.5 hours
Travel time Assume 1 hour per home visit	7 hours	26 hours

Table 8.3 – Time costs per service user associated with training

Training costs per service user	Model developers	Home visitor
Total training time costs per service user, assuming a caseload of 17 individuals and assuming model developers provide training to three home visitors.[35] Model developer: 65.25 hours/17 service users/3 home visitors=1.25 hours per service user for the model developer's time Home visitor: 65.25 hours/17 service users Home visitors=4 hours per service user for the home visitor's time	1.25 hours	4 hours
Total time of professional input	65.25 hours	65.25 hours
5 days (8 hours/day) Group size range: 3 to 4	40 hours	40 hours
9 directly observed field sessions per practitioner (assume 75 min)	11.25 hours	11.25 hours
Travel time per session (1 hour)	14 hours	14 hours

Table 8.4 – Total time costs of the intervention, per service user

Total time costs of the intervention, per service user	Coached	Un-coached
Home visitor – direct provision and training	62.50 hours	62.50 hours
Coached supervisor – direct provision	15.75 hours	Not applicable
Model developer – training	1.25 hours	1.25 hours

[35] In our analysis the cost of training is over-estimated (to some extent) if we assume that the skills will be carried forward with new clients in the second half of the year. However, given the relatively low cost of training, the impact on the results is very marginal.

Child abuse and neglect

Economic Appendix C3.5

Cost–consequence analysis: Trauma-focused CBT (T–CBT) vs. treatment as usual among sexually abused children

Review question 16

What is the impact of social and psychological interventions responding to child sexual abuse?

This report was produced by the Personal Social Services Research Unit at the London School of Economics and Political Science. PSSRU (LSE) is an independent research unit and is contracted as a partner of the NICE Collaborating Centre for Social Care (NCCSC) to carry out the economic reviews of evidence and analyses.

Authors of the report:

Marija Trachtenberg and Jennifer Beecham

Contents

1 Background and aims

The aim of this report is to undertake a cost-consequence analysis based on the intervention described in guideline recommendation 1.7.14, which says to 'consider providing trauma-focused CBT to children (boys and girls) who have been sexually abused. Consider involving the non-abusing parent or carer through either joint or parallel sessions'.[36] The recommendation was developed prior to the economic analysis. This intervention was selected for further economic analysis as there was a lack of cost-effectiveness information and it was considered possible to undertake further economic analysis, as agreed with the Guideline Committee.

The recommendation is based on the results of a good quality meta-analysis, drawing on evidence from 10 non-UK RCTs (9 US and 1 Canadian) with a large combined sample size (n=847) (MacDonald et al 2012). Included studies were conducted before 2001. The review authors find the results of their meta-analysis consistent with past reviews but caution that due to reporting standards and study design, they express more caution about the strength of the findings than have done past reviews.[37]

The intervention is trauma-focused cognitive behavioural therapy (T-CBT) and this is compared to 'treatment as usual'. In 9 studies, treatment as usual was supportive unstructured psychotherapy and one study used a waitlist control group. There were 6 types of T-CBT interventions in the meta-analysis and they varied with respect to the number and duration of sessions, whether they were provided to the child alone (in either group or individual sessions) and whether they were provided to both parent and child (either through separate or joint sessions, in individual or group settings). This is described in more detail further on in this report.

The meta-analysis reports the results for the periods of post-treatment, 3–6 months follow-up, and 1+ years follow-up. The meta-analysis synthesised the results for the primary outcomes of the child's psychological functioning (PTSD, depression, and anxiety) and the child's behaviour problems (externalising behaviour and sexual behaviour).

Participants in the RCTs were recruited from a range of sources. Participants could have been referred to the study by their parents, child protective services, criminal justice system, health and mental health providers (Burke et

[36] The recommendation is based on evidence statements 64–69.

[37] For example, a significant limitation is that 3 of the 10 studies undertook intention-to-treat analysis with the remaining studies only reporting on treatment completers. There is a possibility of effects being biased, but the authors are not sure in which direction (MacDonald et al 2012, p.25). Another issue is 'an absence of strong reporting norms so there remains a risk that the scales reported are a biased representation of those collected by the study authors' (MacDonald et al 2012, p.25). This seems particularly relevant for the outcomes of PTSD and anxiety, where almost every study used a different measurement instrument. This was not the case for depression, child behaviour problems, and child sexual behaviour problems: all studies used the same measurement instrument.

al 1996; Celano 1996; Cohen 1998, 2004; Deblinger 1996, 1999; Deblinger 2001; King 2000)

The indicated population are children and adolescents aged between 2 and 18 years who have experienced sexual abuse (as defined by trialists) and may or may not be symptomatic (for either psychological or behavioural problems) (MacDonald et al 2012, p.17, 24). Five of the 10 studies included a mix of both symptomatic and asymptomatic children and the remaining studies included only symptomatic children. This is important because the authors assume this could limit the ability of the intervention to show effectiveness (MacDonald et al 2012, p. 24). It is also important to note that participants' age varied widely. This may impact the size of the intervention's effectiveness if we believe that age may have a differential impact on different outcomes. We may think this is the case if we consider that there are different patterns of impact for different age groups, especially when considering a developmental perspective. For example, preschool children are likely to experience anxiety, nightmares, externalising behaviour and inappropriate sexual behaviours; school-aged children are likely to experience problems at school, hyperactivity and nightmares, and finally, adolescents more likely to experience depression, generalised anxiety, suicidal, self-injurious behaviour, or substance misuse (MacDonald 2012, p.11).

Is the evidence generalisable to the English context?

A majority of the studies in the meta-analysis provided supportive unstructured psychotherapy as 'treatment-as-usual'. We wanted to find out whether this was true in the English context. Based on our brief review of the literature, it is not clear, on a national level, what services are being provided to sexually abused children (Allnock 2009, p.23). This is partly because children and adolescent mental health services (CAMHS) do not record information as to what proportion of service users are seen as a result of sexual abuse (Allnock 2009, p.23, 64, and p.50 citing Barnes et al 2007). A commissioned review of London's sexual assault services indicated "a lack of appropriate psychological service provision, a lack of service flexibility and choice for follow-up care, and there was little support available for patients, caregivers and families" (RCPCH 2015 p.5, citing NHS England and Kings College Hospital Foundation Trust 2015). What little information we do have is based on a 2008 survey of 21 service managers in the UK. The types of therapeutic services on offer included creative therapies (usually art and play) (63%); counselling, cognitive behavioural therapy and 'other' models (59%); psychodynamic and family therapy (slightly fewer than 40% of services); attachment therapy, narrative therapy, transactional analysis, and sensory motor therapy (less than 25%) (Allnock 2009, p.84). In approximately 33% of services, group work was offered (Allnock 2009, p.84).

In conclusion, it is not clear whether unstructured psychotherapy is 'treatment-as-usual' in the English setting. The implication is that, if less effective treatments are being offered in England, then the treatment effects observed in the meta-analysis are likely to be greater.

2 Methods

2.1 Perspective of the analysis

The aim of this report is to undertake a cost-consequence analysis of T-CBT compared to treatment-as-usual.

The cost perspective includes only intervention costs, which accrue to the NHS or personal social care services sector, depending on the funding arrangement. The time horizon is the same as in the meta-analysis. For intervention costs, this is between 6 to 20 weeks. For outcomes, the time horizon reflects post-treatment, short-term (3–6 months), and long-term (1+ years).

The cost–consequence analysis presents results for the primary outcomes of PTSD, depression, anxiety, sexualised behaviour and externalising behaviour.

2.2 Planned vs. actual analysis on outcomes

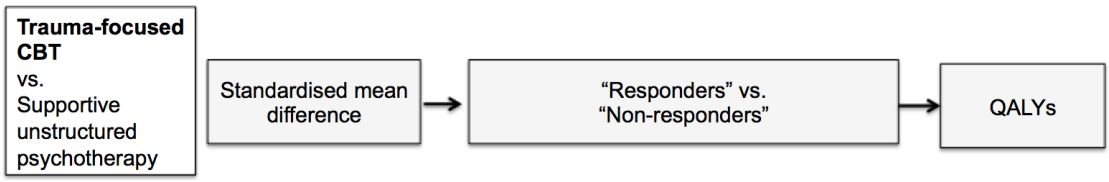
We first present the results as measured by the standardised mean difference (SMD) in effect. The SMD is a way to calculate results for many studies when the studies use different measurement instruments. It is calculated by transforming the results into a uniform scale (Higgings and Green, 2011, section 9.2.3.2). It is calculated as the difference in the mean outcome between groups (i.e. intervention vs. comparison groups) divided by the pooled standard deviation of the outcome of both intervention and control participants (Higgings and Green, 2011) (see equation 1).

Equation 1

$$\text{SMD} = \frac{\text{difference in the mean outcome beteween groups}}{\text{pooled standard deviation of the outcome among all participants}}$$

We planned to then make links from the SMD to changes in quality-adjusted life years (QALYs) (Figure 1). This is only possible if 2 conditions are met. First, we can identify the clinically significant thresholds for the outcome measures. Second, whether there are QALYs that correspond to the clinical thresholds. For example, we would first convert changes in the SMD of depression to the clinical thresholds for ‘treatment responders vs. treatment non-responders’ (i.e. no depression, mild, moderate, or severe). Then we would look for QALYs that correspond to those levels of depression.

Figure 1 – ***Planned analysis on outcomes***



In summary, we found that we could not link any of the outcomes to QALYs.

The first reason is that, for most outcome measures, we could not identify the clinical thresholds. For PTSD (all studies) and some measures of anxiety (3 of 5 studies), the meta-analysis synthesised results for a specific subscale (measurement instruments are provided in Table 1) and there are no clinical thresholds for those subscales. For none of the measures of anxiety (remaining 2/5 studies) and child sexual behaviour (all studies) were we able to make links to clinical thresholds because thresholds are 'normalised' and there are different thresholds depending on different individual characteristics (i.e., child age, gender) and we did not have those normalised thresholds.

The second reason is that, for some outcome measures, while we could identify clinical thresholds, we could not make links to QALYs. For depression (all studies), as measured by the Child Depression Inventory (CDI), it was possible to identify clinical thresholds but it was not possible to make links to QALYs.[38, 39] Whilst clinical thresholds exist, it was not clear what QALY values would be equivalent to 'possible depression' and what value would be used for those who are no longer classified as having 'possible depression', especially considering that they may have comorbidities. We searched Dakin's (2013) database, which provides formulas for linking one outcome measure to various QALY measures. However, no formulas existed to link the CDI to any QALY measurement instrument.

Likewise, for child behaviour problems, as measured by the Child Behaviour Checklist (CBCL), it was possible to make links to clinical thresholds but not to QALYs.[40, 41] Nor did Dakin's (2013) database show a formula linking the CBCL to any QALY measures.

---

[38] Scores for the CDI range from 0 to 54. For scores between 0-13 this indicates no depressive disorder; for scores between 14-19, this indicates possible depressive disorder, scores between 20-54 indicates depressive disorder (Foa et al 2013, p.2652, citing #18).

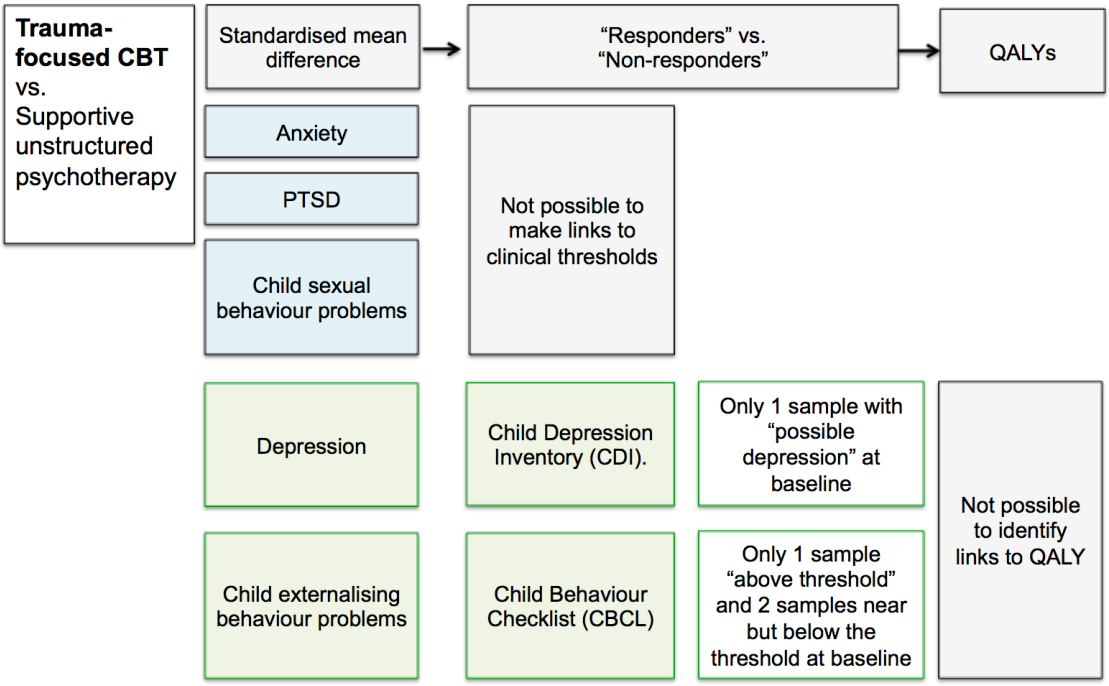
[39] It is also worth noting that, of the five studies available to us in full text, four studies included sample members where the average score indicates that most participants were not depressed at baseline (Berliner 1996, Cohen 1998, Cohen 2004, Deblinger 1999). In the fifth study, participants had a diagnosis of 'possible' depressive disorder at baseline (King 2000). In two remaining studies we could not determine baseline levels of depression (this is because these two papers were PhD theses and we were unable to obtain full text copies (Burke 1998 and Dominguez 2001).

[40] A score of 60+ indicates the child has externalising behaviour problems above the clinical threshold (Newton et al 2000, p. 1366).

[41] Only one study included participants who had clinical thresholds of serious child behaviour problems (King 2000). In another two studies, the participants had baseline scores that were

In conclusion, our cost–consequence analysis takes two approaches (Figure 2). For the outcomes of PTSD, anxiety, and child sexual behaviour, we present results using the standardised mean difference. For the outcomes of depression and externalising behaviour we present the results using both the standardised mean difference and the percentage of individuals who moved between clinical thresholds (reduced symptoms).

Figure 2 – Actual analysis on outcomes



3 Results

3.1 Effectiveness results

The results using the standardised mean difference are presented in Table 1 for PTSD, anxiety, depression, externalising behaviour, and sexual behaviour.

just below clinical thresholds (Cohen 1998; Celano 1996). In the remaining four studies, participants' baseline scores were very low and would not be considered to have clinical levels of externalising behaviour problems (Berliner 1996; Deblinger 1996; Deblinger 1999; Cohen 2004).

Table 1 – Intervention effect size

Note: *Full text was not available ** Effect size is statistically significant

Outcome	Results	Measurement tools
PTSD **Small to medium effect size (p76)	Post-treatment 6 studies, n=442 SMD, -0.44 [-0.73, -0.16]	CITES-R (child report) Celano 1996
	3–6 months after treatment 5 studies, n=327 SMD, -0.39 [-0.74, -0.04]	K-SADS-E, PTSD subscale Deblinger 1996
	1+ years post-treatment 3 studies, n=246 SMD, -0.38 [-0.65, -0.11]	K-SADS-PL (re-experiencing) Cohen 2004
		TSCC PTSD (child report) Cohen 1996, 1998
		ADIS-DSM IV PTSD (re-experiencing) King 2000
Anxiety **Small effect in post-treatment, short-term & long-term (p76)	Post-treatment 5 studies, n=434 SMD, -0.23 [-0.42, -0.03]	Revised Children's Manifest Anxiety Scale (RCMAS) (total) Berliner 1996 King 2000 STAIC (state anxiety) Cohen 1996, 1998 Deblinger 1996 Cohen 2004
	3–6 months after treatment 4 studies, n=296 SMD, -0.38 [-0.61, -0.14]	
	1+ years post-treatment 4 studies, n=278 SMD, -0.28 [-0.52, -0.04]	
Depression Large effect for post-treatment, **short-term & long-term	Post-treatment 5 studies, n=421 SMD, -1.92 [-4.24, 0.40]	Child Depression Inventory *Burke 1988; Berliner 1996 Cohen 1996, 1998 Deblinger 1996 King 2000 *Dominguez 2001 Cohen 2004
	3–6 months after treatment 4 studies, n=286 SMD, -1.84 [-3.41, -0.27]	
	1+ years post-treatment 4 studies, n=301 SMD, -1.19 [-2.70, 0.32]	
Child externalising behaviour problems Equivocal impact. Very small effect sizes in post-treatment, short-term, & long-term.	Post-treatment 7 studies, n=537 SMD, -0.12 [-0.40, 0.17]	Child behavior checklist (CBCL) (externalizing behavior subscale, parent report) Berliner 1996 Celano 1996 Cohen 1996, 1998 Cohen 2004 Deblinger 1996 Deblinger 2001 King 2000
	3-6 months after treatment 4 studies, N=175 SMD, -0.11 [-0.42, 0.21]	
	1+ years post-treatment 5 studies, n=355 SMD, 0.05 [-0.16, 0.27]	
Child sexual behaviour Mixed effects.	Post-treatment 5 studies, n=451 SMD, -0.65 [-3.53, 2.24]	Children's Sexual Behavior Inventory (CSBI) Berliner 1996

Medium effect post-treatment. Small/medium effect in short-term. Large effect in long-term	3-6 months after treatment 3 studies, n=133 SMD, -0.46 [-5.68, 4.76]	Cohen 1996, 1998 Deblinger 2001 Cohen 2004
	1+ years post-treatment 3 studies, n=161 SMD, -1.61 [-5.72, 2.49]	

Table 2 and Table 3 (below) present the results for depression and externalising behaviour as the percentage of individuals improving, as determined by the clinical thresholds.

We calculated this percentage using a Monte Carlo simulation using several steps. The calculations are based on a hypothetically large sample (which we chose to be 1,000, as is standard practice). We then created a 'randomisation' formula that picks a value within the constraints of the mean and standard error for each outcome (i.e. depression or externalising behaviour). We assumed that the chances of getting different numbers are based on a 'normal distribution' (i.e. that most of the values are near the mean and that the values that are further away from the mean have a smaller chance of occurring, and that the chances are evenly distributed on both sides of the mean).

We created tables that contained the baseline scores (simulated 1,000 times) and the post-treatment scores (baseline scores + standardised mean difference). A third table was created to indicate whether the individual is above or below the clinical threshold. If the baseline score was below the clinical threshold, then the result was 'not applicable'. If the baseline score was above the clinical threshold and the post-treatment score was below the clinical threshold, we counted that individual as 'being below the threshold'. If the individual remained above the threshold, that individual was counted as 'remaining above the threshold'. The final step was to calculate the percentage of individuals that improved based on the 1,000 simulations. We repeated this process for each time period: post-treatment, 3–6 months follow-up, and 1+ years follow-up. These calculations were completed in MS Excel.

For the outcome of depression, the simulations show that between 0% and 10.8% of individuals improved in the post-treatment period, between 0% and 12.1% improved in the 3–6 month follow-up period, and between 0% and 10.7% improved in the 1+ year follow-up period. These results are dependent on both the initial baseline depression scores and the size of the intervention's effect on reducing depressive symptoms. It is important to note that the size of the intervention's effect was large, but that this was only statistically significant in the short-term (3–6 months follow-up). This is because in the short-term, the mean reduction in depression scores was -1.84 and all of the values in the 95% confidence interval were also showing reductions [95% CI, -3.41, -0.27]. This is in contrast to the post-treatment and 1+ year follow-up effect, which found that while many individuals had reduced depressive symptoms, some

individuals did not do better than the comparison group (as indicated by the 95% confidence interval having both negative and positive scores).[42]

Table 2 – Depression: percentage of individuals with improved symptoms (falling below clinical threshold at different time periods) based on 1,000 simulations

	Berliner 1996	Cohen 1996, 1998	Cohen 2004	Deblinger 1999	King 2000
Post-treatment	0.2%	8.7%	0%	2.5%	10.8%
3–6 month follow-up	0.2%	5.1%	0%	2.7%	12.1%
1+ years follow-up	0.2%	1.8%	0%	2.0%	10.7%

For the outcome of externalising behaviour, the meta-analysis showed that the intervention had very small impact and this was not statistically significant, which indicates that, on the whole, the intervention may not be any better at reducing externalising behaviour compared to unstructured psychotherapy. This is because while the intervention did result in some individuals having reduced externalising behaviour problems, there are still many individuals who did not do better (as indicated by the 95% confidence interval having both negative and positive scores).[43] The results from the simulation indicate that, in combination with baseline scores and size of the intervention effect, none of the individuals moved below the clinical threshold for all time periods.

Table 3 – Externalising behaviour: percentage of individuals with improved symptoms (falling below clinical threshold at different time periods) based on 1,000 simulations

	Berliner 1996	Celano 1996	Cohen 1996, 98	Cohen 2004	Deblinger 1996, 99	Deblinger 2001	King 2000
Post-treatment	0%	0%	0%	0%	0%	0%	0%
3-6 months follow-up	0%	0%	0%	0%	0%	0%	0%
1+	0%	0%	0%	0%	0%	0%	0%

[42] The mean reduction and 95% confidence intervals for depressive symptoms for the post-treatment and 1+ year follow-up periods are, respectively, -1.92 [95% CI, -4.24, 0.40] and -1.19 [95% CI, -2.70, 0.32].

[43] The mean and 95% confidence intervals for externalising behaviour problems are, for post-treatment, SMD, -0.12 [95% CI, -0.40, 0.17], 3–6 month follow-up, SMD, -0.11 [95% CI, -0.42, 0.21], and 1+ year follow-up, SMD, 0.05 [95% CI, -0.16, 0.27].

years
follow-up

3.2 Intervention costs

Our costs are based on national average estimates and use a full-cost approach in line with accepted practice (Curtis and Burns 2016). All costs reflect 2015/16 price year.

The costs of the intervention vary because the meta-analysis combined several different types of T-CBT interventions, including:

1. Child only group
2. Child only group + parent support group (not CBT)
3. Child only group + parent only group
4. Individual sessions for the child
5. Individual sessions for the child and parent (joint sessions)
6. Individual sessions for the child and parent (separate sessions)

We provide estimates of intervention costs for each type of T-CBT included in the meta-analysis.[44] The studies did not always provide all the information needed to estimate costs. In these instances we made assumptions based on information in other studies.

There are three main intervention costs. The first is the therapist time for directly providing the intervention (and the associated time to complete administrative, patient-related tasks). The second is the time to train the therapist. The third is supervision, time required of the supervisor and the therapist.

We provide two sets of intervention costs. One set of costs includes only the therapist's time in providing the intervention and the time for training; effectively, it excludes the costs of supervision. The second set of costs includes all three cost components (i.e. includes supervision costs). The reason we present two sets of costs is that clinical research studies may have provided supervision to ensure that the therapists are delivering the intervention 'as intended' (i.e. fidelity to the model). In real world situations, supervision may not occur as frequently or may not be provided at all.

3.3 Intervention costs influenced by intervention characteristics

We also present intervention costs using lower, middle and upper estimates, which are influenced by the intervention characteristics: (a) size of the group session (where appropriate), (b) number of sessions, (c) session duration, (d) time spent on patient-related administration, (e) number of therapists per session (applicable to group sessions), (f) hours of training for therapists, (g) hours of weekly supervision, (h) whether supervision was provided 1:1 or to several therapists, (i) time required for the supervisor to review audio-taped sessions on a weekly basis.

[44] We could not estimate costs for 2 studies because these were PhD theses and we could not obtain these in full-text (Dominguez 2001 and Cohen 1996). Studies that were included can be found in the appendix.

Intervention inputs are described in Table 4. Our estimates and assumptions used for intervention costing are provided in Table 5 (total therapist time for intervention delivery) and Table 6 (total time required of therapists and supervisors for training and supervision).

Supervision

In 5 studies, supervision was clearly stated, which involved weekly supervision with or without the supervisor reviewing audiotaped sessions each week.[45] However, none of these studies specified the duration of supervision so we used assumptions.[46] In most studies it was also not clear whether supervisors conducted 1:1 weekly supervisions with one therapist or with several therapists. We assumed a ratio of 1 supervisor to 2 therapists per week.[47] Our cost estimates relating to supervision also assumed that supervisors review audiotaped sessions each week, and that they listened to the entire audiotape. This effectively almost doubled the intervention costs. This was done in three studies and not mentioned in the remaining studies.[48]

Training

Only 3 studies provided information about training and duration.[49] Training was not described in the remaining studies. We assumed 15 hours of training per therapist. We assumed that the per-person costs of training could be shared amongst a hypothetical group of 20 recipients, which is a conservative estimate. It is possible that training could last for more than 20 individuals.[50]

Intervention format and delivery

Three interventions were delivered in a group format but none of the studies specified group size. We assumed group sizes of three, five, and seven, based on similar group sizes in related research.

The number and duration of sessions varied. Group-based sessions ranged from 6 to 11 sessions with a duration ranging from 60 minutes to 1 hour and 45 minutes for each session. In the individual-only sessions, there were between 12 and 20 sessions and duration ranged from 45 to 90 minutes.

---

[45] Supervision was clearly stated for the following 5 studies: Celano 1996, Cohen 1998, Cohen 2004, Deblinger 1996 and 1999; King 2000. It was not clear whether supervision was provided in the remaining studies.

[46] We assumed that supervision lasted 1 hour in most scenarios. However, we undertook scenario analyses where we assumed 2-hour supervision scenarios where there was a group size of 7. In a group size of 5, we assumed supervision lasted 1.5 hours. For group sizes of 3 and in all individual sessions, we assumed supervision lasted 1 hour.

[47] In the calculations, this is represented as a ratio of 0.5 (1 supervisor to 2 therapists).

[48] There were 3 studies where supervisors reviewed therapists' audiotapes each week: Cohen 1998, Cohen 2004, Deblinger 1996 and 1999.

[49] Three studies described duration of training: Celano 1996; Cohen 2004; King 2000.

[50] Per person, this is an additional cost 0.8 hours (for group sizes 5 and less, provided by one therapist) and an additional cost of 1.5 hours (for a group size of 7, provided by two therapists).

In all of the individual-only sessions, only 1 therapist provided the intervention. In the group-based sessions, most studies did not specify, with the exception of 1 study where 2 therapists provided the intervention. We estimated the cost of the group-based interventions assuming that 2 therapists provide the intervention where group size is 7 and only 1 therapist is involved where the group size is 5 or less.

We also included the cost of therapists' time for patient-related administrative activity that might be done after the session. In the absence of data from the studies, we assumed an additional 15 minutes of patient-related work per person.[51]

[51] In the calculations, the ratio used in the calculation is 1.25, representing 15 minutes of additional patient-related work.

Table 4 – Intervention inputs: delivery, training, and supervision

	Group size	Sessions	Session duration	Total therapist hours for all sessions	Number of therapists per session	Who delivers intervention	Training	Supervision	Super-vision duration
(1) Child only group									not reported
Burke 1988	not reported	6	not reported	not reported	not reported	not reported	not reported	not reported	
Celano 1996	not reported	8	60	8	not reported	Psychiatrists, psychologists, social workers, nurses, & trainees in psychiatry & psychology	3 hrs	(1) Weekly supervision (1 professional to 3 trainees)	
(2) Child only group + parent support group (not CBT)									
Berliner 1996	not reported	10	not reported	not reported	2	Masters level clinical social worker	not reported	not reported	
(3) Child only group + parent only group									
Deblinger 2001	not reported	11	105 min/ session + 15 min joint session	22	1	not reported	not reported	not reported	
(4) Individual sessions for the child									
King 2000	n/a	20	50	16.7	1	not reported	15-20 hrs	(1) Weekly supervision	
Deblinger 1996, 1999	n/a	12	45	9	1	Mental health therapist	not reported	(1) Weekly supervision by PI (2) supervisors listen to audiotapes weekly	
(5) Individual sessions for the child and parent (joint sessions)									
Deblinger 1996, 1999	n/a	12	90	17	1	Mental health therapist	not reported	(1) Weekly supervision (2) Supervisors listen to audiotapes weekly	
(6) Individual sessions for the child and parent (separate sessions)									
Cohen 1998	n/a	12	45	18	1	Masters level clinical social worker	not reported	(1) Weekly supervision + (2) weekly audio review	
Cohen 2004	n/a	12 individual + 3 joint	45 individual + 30 min joint	19.5	1	Psychologist / Social worker	3 days	(1) Weekly (2) review of all audiotapes (3) 2x monthly cross-site phone call	
King 2000	n/a	20	50	16.7	1	not reported	15-20 hrs	(1) Weekly supervision	
Unclear									
Cohen 1996 and Dominguez 2001. It was not possible to obtain full text for both of these papers (both are PhD theses).									

Table 5 – Intervention costing: total therapist time for intervention delivery

	Group size (a)	Sessions (b)	Session duration (c)	Ratio of direct to indirect time with client (multiplier) (d)	Number of therapists per session (e)	Total hours for all sessions for all therapists involved per treatment unit (i.e. individual or group) (f) = [(b*c*e)+(c*d)]/ 60 min	Total therapist time for intervention delivery, per person (g) = (f/a)
(1) Child only group							
Lower estimate	7	6	60	1.25	2	15.0	2.1
Middle estimate	5	7	60	1.25	1	8.8	1.8
Upper estimate	3	8	60	1.25	1	10.0	3.3
(2) Child only group + parent support group (not CBT)							
Lower estimate	7	10	60	1.25	2	25.0	3.6
Upper estimate	3	10	60	1.25	1	12.5	4.2
(3) Child only group + parent only group							
Lower estimate	7	10	225	1.25	1	46.9	6.7
Middle estimate	5	10	225	1.25	1	46.9	9.4
Upper estimate	3	11	225	1.25	1	51.6	17.2
(4) Individual sessions for the child							
Lower estimate	1	12	45	1.25	1	11.3	11.3
Middle estimate	1	16	47.5	1.25	1	15.8	15.8
Upper estimate	1	20	50	1.25	1	20.8	20.8
(5) Individual sessions for the child and parent (joint sessions)							
Lower estimate	1	12	90	1.25	1	22.5	22.5
Middle estimate	1	12	90	1.25	1	22.5	22.5
Upper estimate	1	12	90	1.25	1	22.5	22.5
(6) Individual sessions for the child and parent (separate sessions)							
Lower estimate	1	12	45	1.25	1	22.5	22.5
Middle estimate	1	16	47.5	1.25	1	31.7	31.7
Upper estimate	1	20	50	1.25	1	41.7	41.7

Table 6 – Intervention costing: training and supervision

	Total therapist training hours * number of therapists per session (h)	Total therapist training (hours per person) (i) = (h/20)	Weekly supervision ratio (supervisor to therapist) (k)	Supervision per week (hours) (m)	Supervisor audiotape review, duration (hours) (n)	Total supervisor time for supervision, hours per person (p) = [(n)+((m*b)*k)]/a	Total therapist supervision time, hours per person (q) = [(n +(m*b)]/a
(1) Child only group							
Lower estimate	30	1.5	0.5	2	15.0	1.9	3.9
Middle estimate	15	0.8	0.5	1.5	8.8	1.9	3.9
Upper estimate	15	0.8	0.5	1	10.0	3.0	6.0
(2) Child only group + parent support group (not CBT)							
Lower estimate	30	1.5	0.5	2	25.0	3.2	6.4
Upper estimate	15	0.8	0.5	1	12.5	3.8	7.5
(3) Child only group + parent only group							
Lower estimate	15	0.8	0.5	2	46.9	4.8	9.6
Middle estimate	15	0.8	0.5	1.5	46.9	6.2	12.4
Upper estimate	15	0.8	0.5	1	51.6	10.4	20.9
(4) Individual sessions for the child							
Lower estimate	15	0.8	0.5	1	11.3	11.6	23.3
Middle estimate	15	0.8	0.5	1	15.8	15.9	31.8
Upper estimate	15	0.8	0.5	1	20.8	20.4	40.8
(5) Individual sessions for the child and parent (joint sessions)							
Low estimate	15	0.8	0.5	1	22.5	17.3	34.5
Middle estimate	15	0.8	0.5	1	22.5	17.3	34.5
High estimate	15	0.8	0.5	1	22.5	17.3	34.5
(6) Individual sessions for the child and parent (separate sessions)							
Lower estimate	15	0.8	0.5	1	22.5	17.3	34.5
Middle estimate	15	0.8	0.5	1	31.7	23.8	47.7
Upper estimate	15	0.8	0.5	1	41.7	30.8	61.7

3.4 Intervention costs influenced by who delivers the intervention

We also present lower, middle, and upper intervention cost estimates depending on *the professional* providing the intervention. The types of providing professionals varied.[52] We estimated UK-equivalents to be a mental health nurse, children's social worker and a clinical psychologist or consultant psychiatrist. The hourly unit cost for each professional's time is, respectively, £40/hour, £57/hour, and £139/hour (Curtis and Burns 2015).

3.5 Intervention cost estimates

The intervention cost estimates are provided in Tables 7 and 8. Cost estimates are presented per 'treatment unit'. This means that if the sessions are provided for the child only, then the costs relate to the child. Where the sessions are provided to both parent and child (jointly or separately), 'treatment unit' reflects the cost of providing therapy to the parent-child dyad. Table 7 provides the cost estimates (per treatment unit) where supervision is excluded. Table 8 provides the cost estimates (per treatment unit) where supervision costs are included. Across intervention types, the inclusion of supervision triples the intervention cost. As expected, group-based T-CBT costs less than individual T-CBT. Likewise, interventions provided by a clinical psychologist or consultant psychiatrist are more costly than when provided by a mental health nurse or children's social worker.

Intervention costs per treatment unit vary (rounded to nearest hundred):

1. Child group sessions range from £100 to £570 when excluding supervision costs and increase to a range between £300 and £1,800 when including supervision costs.

2. Child group sessions + parent support group (not CBT) range from £200 to £700 (excluding supervision) and £600 to £2,200 (including supervision).

3. Child group sessions + parent group sessions (both are T-CBT) range from £300 to £2,500 (excluding supervision) and £900 to £6,800 (including supervision).

4. Individual child sessions range from £500 to £3,000 (excluding supervision) and £1,800 to £11,500 (including supervision).

5. Joint individual sessions for child and parent range from £900 to £3,200 (excluding supervision) and increase to £3,000 to £10,400 (including supervision).

6. Separate individual sessions for the child and parent (separate sessions) range from £900 to £5,900 (excluding supervision) and increase to £3,000 and £18,700 (including supervision).

[52] Masters level clinical social worker (Berliner 1996; Cohen 1998), social worker (Cohen 2004), psychologist (Celano 1996, Cohen 2004), mental health therapist (Deblinger 1996, 1999), psychiatrist or nurse (Celano 1996), "therapist" unspecified (Burke 1998, Deblinger 2001, King 2000).

Table 7 – Intervention costing: total cost per treatment unit, EXCLUDING SUPERVISION

2015/16 prices	**Total cost per treatment unit = (g + i) * unit cost/hour**		
	Lower estimate Mental health nurse (£40/hour)	**Middle estimate** Children's social worker (£57/hour)	**Upper estimate** Psychiatrist or clinical psychologist (£139/hour)
(1) Child only group			
Lower estimate	£150	£210	£510
Middle estimate	£100	£140	£350
Upper estimate	£160	£230	£570
(2) Child only group + parent support group (not CBT)			
Lower estimate	£200	£300	£700
Upper estimate	£200	£300	£700
(3) Child only group + parent only group			
Lower estimate	£300	£400	£1,000
Middle estimate	£400	£600	£1,400
Upper estimate	£700	£1,000	£2,500
(4) Individual sessions for the child			
Lower estimate	£500	£700	£1,700
Middle estimate	£700	£900	£2,300
Upper estimate	£900	£1,200	£3,000
(5) Individual sessions for the child and parent (joint sessions)			
Lower estimate	£900	£1,300	£3,200
Middle estimate	£900	£1,300	£3,200
Upper estimate	£900	£1,300	£3,200
(6) Individual sessions for the child and parent (separate sessions)			
Lower estimate	£900	£1,300	£3,200
Middle estimate	£1,300	£1,800	£4,500
Upper estimate	£1,700	£2,400	£5,900
NOTE: Figures are rounded to nearest 100			

Table 8 – Intervention costing: total cost per treatment unit, INCLUDING SUPERVISION

2015/16 prices	**Total cost per treatment unit = (g + i + p + q) * unit cost/hour**		
	Lower estimate Mental health nurse (£40/hour)	**Middle estimate** Children's social worker (£57/hour)	**Upper estimate** Psychiatrist or clinical psychologist (£139/hour)
(1) Child only group			
Lower estimate	£400	£500	£1,300
Middle estimate	£300	£500	£1,200
Upper estimate	£500	£700	£1,800
(2) Child only group + Parent support group (not CBT)			
Lower estimate	£600	£800	£2,000
Upper estimate	£600	£900	£2,200
(3) Child only group + Parent only group			
Lower estimate	£900	£1,200	£3,000
Middle estimate	£1,100	£1,600	£4,000
Upper estimate	£2,000	£2,900	£6,800
(4) Individual sessions for the child			
Lower estimate	£1,800	£2,700	£6,500
Middle estimate	£2,600	£3,700	£8,900
Upper estimate	£3,300	£4,700	£11,500
(5) Individual sessions for the child and parent (joint sessions)			
Lower estimate	£3,000	£4,300	£10,400
Middle estimate	£3,000	£4,300	£10,400
Upper estimate	£3,000	£4,300	£10,400
(6) Individual sessions for the child and parent (separate sessions)			
Lower estimate	£3,000	£4,300	£10,400
Middle estimate	£4,200	£5,900	£14,400
Upper estimate	£5,400	£7,700	£18,700
NOTE: Figures are rounded to nearest 100			

4 Discussion

4.1 Conclusions about cost-effectiveness

In summary, T-CBT is more effective than supportive unstructured psychotherapy for the outcomes of PTSD and anxiety and has weak evidence of effectiveness for depression. For the outcomes of child behaviour problems and sexual behaviour, T-CBT was not more effective than supportive unstructured psychotherapy. The additional costs of T-CBT to achieve those outcomes vary across the six types of T-CBT interventions described previously.

There is not enough evidence to assess the cost-effectiveness of T-CBT compared to supportive unstructured psychotherapy.

This analysis provided information about the additional costs of the intervention and how the intervention changes outcomes for children. In this sense, our cost-consequence analysis is missing additional information, such as the potential impact of the intervention on an individual's use of public sector services, which would give this economic analysis a wider perspective. There was no information about wider impacts on service use. In conclusion, decision makers will need to decide whether the additional improvements for children are worth the additional costs.

Table 9 – Costs and effects of T-CBT vs. supportive unstructured psychotherapy

Additional costs of different types of T-CBT			**Effectiveness of T-CBT**
Type	Including supervision	Excluding supervision	
1	£300 to £1,800	£100 to £570	Reduced PTSD symptoms
2	£600 to £2,200	£200 to £700	Reduced anxiety symptoms
3	£900 to £6,800	£300 to £2,500	Weak reduction in depressive symptoms
4	£1,800 to £11,500	£500 to £3,000	No differences in child behaviour problems
5	£3,000 to £10,400	£900 to £3,200	No differences in sexual behaviour
6	£3,000 and £18,700	£900 to £5,900	

If decision makers decide that T-CBT is cost-effective, it is important to note that we cannot conclude that the 'less costly' group-based T-CBT is more cost-effective than the 'more costly', individual T-CBT. Likewise, we cannot conclude that the 'less costly' provision of T-CBT by mental health nurse or child social worker is more cost-effective compared to the 'more costly' clinical psychologist and consultant psychiatrist. The reason is that these studies

were not designed to answer those questions. Rather, they were designed to conclude whether T-CBT is more effective than supportive unstructured psychotherapy. Moreover, the meta-analysis combined the available studies even though T-CBT was provided in different ways and delivered by different professionals.

4.2 Implications for decision-makers

The meta-analysis shows that the intervention is effective for reducing PTSD symptoms, with a small to medium effect size that is statistically significant, across post-treatment to 1+ years follow-up (MacDonald 2012, p.76). This is important because PTSD is the most frequently observed symptom among sexually abused children and children are usually comorbid with depression and anxiety (Gospodarevskaya 2012, p.1), for which this meta-analysis found that T-CBT was effective for improving symptoms of anxiety, with a small effect size that is statistically significant across post-treatment to 1+ years follow-up (MacDonald 2012, p.76). Likewise, the meta-analysis found T-CBT to be effective in reducing symptoms of depression, with a large effect size in the post-treatment and 1+ years follow-up period (but not statistically significant), and a large effect size that is statistically significant in the 3-6 month follow-up period (MacDonald 2012, p.76).

While we could not make links from changes in symptoms to changes in QALYs (due to issues described earlier), the quality of life with PTSD is low. Based on 1 Australian study, children with or without experience of child sexual abuse (before aged 18) had a health state utility of 0.87 (the state of 'best' health is 1.0). The health state utility of having experienced child sexual abuse before aged 18 years is 0.71, regardless of whether they met criteria for PTSD, and 0.61 among those who did meet the criteria for PTSD (Gospodarevskaya 2013, p.279). Among those who met the criteria for both PTSD and depression, the health state utility was much lower, at 0.53 (Gospodarevskaya 2013, p.279). However, these results need to considered with caution as we do not have estimates taken from a UK population and that the Australian study relies on a small sample.[53] While it is not clear how much QALYs would have improved based on T-CBT, we can appreciate that any improvements are significant, especially when they occur across multiple outcomes (PTSD, anxiety, depression).

Furthermore, our additional analysis indicates that the percentage of individuals that no longer have 'possible depression' and are now categorised as 'not depressed' (using the Child Depression Inventory) is between 0% and 10.8% for the post-treatment period, between 0% and 12.1% for the 3-6 month follow-up period, and between 0% and 10.7% for the 1+ year follow-up period. These estimates are calculated using simulation analysis and are dependent on the baseline scores of participants and the effectiveness of T-CBT (as presented earlier in Tables 2 and 3).

[53] Total sample size for the survey was based on n=993. Total number experiencing sexual abuse before aged 18 years is n=82. Total number meeting criteria for PTSD and having experienced sexual abuse before aged 18 years is n=14. Total number meeting criteria for PTSD and depression and having experienced sexual abuse before aged 18 years is n=9.

Given that there are several types of T-CBT interventions, decision makers may want to consider individuals' preferences when deciding which type of intervention to offer. Of course, this will need to be balanced against resource constraints. It may be worthwhile considering partnerships between different local authorities and NHS trusts to share the provision of group-based sessions, given that the number of individuals that have experienced child sexual abuse may be too few to deliver within a single local authority or NHS area. Likewise, where waiting lists exist for individual sessions in one area, partnerships with other areas with spare capacity to cross-deliver services may help alleviate waiting lists.

4.3 Limitations and implications for decision makers and research

Our analysis is limited as we could only consider intervention costs only. There was no information about how T-CBT might change an individuals' current and future use of health and social care services, compared to an individual who receives unstructured psychotherapy. In this sense, our cost-consequence analysis is missing additional information about the potential wider impacts on public sector costs, which would otherwise make the results of this economic analysis much more comprehensive.

Future research should include an economic evaluation alongside an evaluation of an intervention's effectiveness. Future research should also include a long follow-up period to understand whether treatment effects are sustained and how service use changes over time.

5 Linking economic evidence to recommendations

Based on the effectiveness evidence alone, the Guideline Committee had already decided to recommend this intervention prior to the economic analysis.

We noted to the Guideline Committee that they will need to comment on the likelihood of supervision being of equal intensity to what is described in clinical research studies or not (or of lesser intensity) as the intensity of the supervision has a significant effect on intervention costs.

The Guideline Committee still decided it was worth recommending the Trauma-focused CBT intervention based on the results of the study's reported effectiveness. Given the potentially large resource implications of implementing the intervention, NICE requested that scenario analyses be carried out, and to estimate the QALY gain and/or cost offsets required for the Trauma-focused CBT intervention to be cost-effective. After performing scenario analyses, the Guideline Committee believed that the Trauma-focused CBT intervention could be cost-effective at £20,000-£30,000 per QALY, based on the assumptions made in the scenario analyses. The assumption is that Trauma-focused CBT results in QALY gains of between 0.12 and 0.30 over a 1-year period, assuming that either the less- or more-resource intensive version of the intervention is provided. The Guideline Committee assumed that these gains could be generated by improvements to

the child's level of anxiety/depression (a key area of the EQ-5D) based on the study's reported findings of improvements in the child's PTSD and anxiety at post-treatment, 3-6 months follow-up, and 1-year follow-up. This is based on the assumption that the intervention would not lead to changes in public sector service use during that 1-year period. For further detail on the scenario analyses, please refer to NICE to request the documents.

Recommendation 1.7.17 says:

Offer group or individual trauma-focused cognitive behavioural therapy over 12 to 16 sessions (more if needed) to children and young people (boys or girls) who have been sexually abused and show symptoms of anxiety, sexualised behaviour or post-traumatic stress disorder. When offering this therapy:

- discuss it fully with the child or young person before providing it and make clear that there are other options available if they would prefer
- provide separate trauma-focused cognitive behavioural therapy sessions for the non-abusing parent or carer.

6 Bibliography

Allnock D, Bunting L, Price A, Morgan-Klein N, Ellis J, Radford L, Stafford A (2009). Sexual abuse and therapeutic services for children and young people: the gap between provision and need. NSPCC <https://www.nspcc.org.uk/globalassets/documents/research-reports/sexual-abuse-therapeutic-services-children-young-people-report.pdf>

Dakin H (2013). Review of studies mapping from quality of life or clinical measures to EQ-5D: an online database. Health and Quality of Life Outcomes. 11:151. HERC database of mapping studies, Version 3.0 (Last updated: 26th June 2014). Available at: http://www.herc.ox.ac.uk/downloads/mappingdatabase

Department of Health and Department of Education Services (2006). Report on the Implementation of Standard 9 of the National Service Framework for Children, Young People and Maternity Services. London: DH

Gospodarevskaya E (2013). Post-traumatic stress disorder and quality of life in sexually abused Australian children. Journal of child sexual abuse, 22(3): 277–96

Gospodarevskaya, E, Segal L (2012). Cost-utility analysis of different treatments for post-traumatic stress disorder in sexually abused children. Child and Adolescent Psychiatry and Mental Health, 6: 15

Higgings JPT, Green S (2011). Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. Cochrane Collaboration http://handbook.cochrane.org/front_page.htm

Macdonald G, Higgins JPT, Ramchandani P, Valentine JC, Bronger LP, Klein P, O'Daniel R, Pickering M, Rademaker B, Richardson G, Taylor M (2012).

Cognitive-behavioural interventions for children who have been sexually abused. Campbell Systematic Reviews 14

Newton RR, Litrownik AJ, Landsverk JA (2000). Children and youth in foster care: disentangling the relationship between problem behaviours and number of placements. Child abuse and neglect, 24(10): 1363–74

NHS England and Kings College Hospital Foundation Trust (2015). Review of pathway following sexual assault for children and young people in London. March. http://www.england.nhs.uk/london/wp-content/uploads/sites/8/2015/03/review-pathway- cyp-london-summary.pdf

RCPCH (Royal College of Pediatrics and Child Health) (2015). Service specification for the clinical evaluation of children and young people who may have been sexually abused http://www.rcpch.ac.uk/system/files/protected/page/Service%20Specification%20for%20the%20clinical%20evaluation%20of%20CYP%20who%20may%20have%20been%20sexually%20abused_September_2015_FINAL.pdf

Studies included in the meta-analysis

Berliner 1996

Berliner L, Saunders BE (1996). Treating fear and anxiety in sexually abused children: results of a controlled 2-year follow-up study, Child Maltreatment, 1(4): 294–309

Burke 1998

Burke MM (1988). Short-term Group Therapy for Sexually Abused Girls: A Learning Theory Based Treatment for Negative Affect [Phd dissertation]. Athens, Georgia: University of Georgia. Unpublished data only

Celano 1996

Celano M, Hazzard A, Webb C, McCall C (1996). Treatment of traumagenic beliefs among sexually abused girls and their mothers: an evaluation study. Journal of Abnormal Child Psychology, 24(1): 1–17

Cohen 1998, 2005

Cohen JA, Mannarino AP (1998). Interventions for sexually abused children: initial treatment outcome findings. Child Maltreatment, 3(1): 17–26

Cohen JA, Mannarino AP, Knudsen K (2005). Treating sexually abused children: 1 year follow-up of a randomized controlled trial. Child Abuse & Neglect, 29(2): 135–45

Cohen 2004

Cohen JA, Deblinger E, Mannarino AP, Steer RA (2004). A multisite, randomized controlled trial for children with sexual abuse-related PTSD

symptoms. Journal of the American Academy of Child and Adolescent Psychiatry, 43(4): 393–402

Deblinger E, Mannarino AP, Cohen JA, Steer RA (2006). A follow-up study of a multisite, randomized controlled trial for children with sexual abuse-related PTSD symptoms. Journal of the American Academy of Child and Adolescent Psychiatry, 45(12): 1474–84

Deblinger 1996

Deblinger E, Lippman J, Steer R (1996). Sexually abused children suffering post traumatic stress symptoms. Child Maltreatment, 1(4): 310–21

Deblinger E, Steer R, Lippman J (1999). Two-year follow-up study of cognitive behavioral therapy for sexually abused children suffering post-traumatic stress symptoms. Child Abuse and Neglect, 23(12): 1371–8

Deblinger 2001

Deblinger E, Stauffer LA, Steer RA (2001). Comparative efficacies of supportive and cognitive behavioral group therapies for young children who have been sexually abused and their nonoffending mothers. Child Maltreatment, 6(4): 332–43

Dominguez 2001

Dominguez RZ.(2001). Evaluation of Cognitive-Behavioral and Supportive Treatments for Sexually Abused Children: Analyzing the Process of Change Using Individual Growth Curve Analyses [Phd dissertation]. Houston TX: University of Houston

King 2000

King NJ, Tonge BJ, Mullen P, Myerson N, Heyne D, Rollings S et al. (2000). Treating sexually abused children with posttraumatic stress symptoms: a randomized clinical trial. Journal of the American Academy of Child and Adolescent Psychiatry, 39(11): 1347–55

7 Appendix

8.1 Baseline scores of various outcome measures

Baseline PTSD score	Mean	SE
Celano 1996, CITES-R (child report)		
Intervention	46.7	2.40
Comparison	46.2	1.73
Cohen 1998, TSCC (child report) PTSD subscale		
Intervention	10.63	0.81
Comparison	10.83	0.91
Cohen 2004, K-SADS-PL (re-experiencing)		
Intervention, N=89	3.98	0.12
Comparison, N=91	4.08	0.12
Deblinger 1999, K-SADS-E, PTSD subscale		
Intervention	9.7	0.65
Comparison	8.86	0.59
Comparison	10.47	0.61
Comparison	9.79	0.64
King 2000, ADIS-DSM IV PTSD (re-experiencing)		
Intervention	3.33	0.43
Comparison	4.16	0.27
Wait-list control	3.33	0.43

Baseline anxiety score	Mean	SE
Berliner 1996, RCMAS (total score)		
Intervention	14.1	1.36
Control	14.5	1.25
King 2000, RCMAS (total score)		
Intervention	14.1	1.36
Comparison	14.5	1.25
WLC	14.1	1.36
Cohen 1998, STAIC (state anxiety)		
Intervention	35.32	1.51
Comparison	34.45	1.54
Cohen 2004, STAIC (state anxiety)		
Intervention	30.51	0.71
Comparison	31.48	0.87
Deblinger 1999, STAIC (state anxiety)		
Intervention	30.82	1.32
Comparison	33.73	1.56
Comparison	28.42	1.19
Comparison	31.09	1.13

Baseline depression (CDI)	Mean	SE
Berliner 1996		
Intervention	9.7	1.65
Control	10.1	1.38
Cohen 1998		
Intervention	12.37	1.38
Comparison	11.7	1.17
Cohen 2004		
Intervention	9.92	0.78
Comparison	12.11	0.90
Deblinger 1999		
Intervention	12.15	1.38
Comparison	9.81	1.24
Comparison	10.84	1.57
Comparison	11.87	1.30
King 2000		
Intervention	16.83	2.85
Comparison	18.83	3.17
Wait-list control	17.33	2.38
Dominguez 2001	Full-text not available	
Burke 1988		

Baseline externalising behaviour score (CBCL)	Mean	SE
Berliner 1996		
Intervention	15.0	1.36
Control	18.2	1.89
Celano 1996		
Intervention	55.1	1.8
Comparison	66.2	2.6
Cohen 1998		
Intervention	57.61	2.30
Comparison	56.23	2.56
Deblinger 1996		
Intervention	13.27	1.59
Comparison	15.09	2.08
Comparison	19.5	2.04
King 2000		
Intervention	67.17	4.21
Comparison	73.08	3.45
WLC	64.58	4.34
Deblinger 1999		
Intervention	12.44	1.54
Comparison	18.75	2.14
Comparison	15.72	2.37
Comparison	14.92	3.11
Cohen 2004		
Intervention	15.59	1.12
Comparison	17.18	1.04