Appendix N: GRADE evidence profiles for all studies

Abbreviations

AAMD	American Association on Mental Deficiency
ADHD	attention deficit hyperactivity disorder
A-PS	assertiveness then social problem-solving
BDI(-II)	Beck Depression Inventory (revised)
CBT	cognitive behavioural therapy
CES-D	Center for Epidemiologic Studies - Depression scale
CGI	Clinical Global Impression scale
CI	confidence interval
GRADE	Grades of Recommendation Assessment, Development and Evaluation
GSI	Global Severity Index
HAM-A	Hamilton Anxiety Rating Scale
ITT	intention to treat
MD	mean difference
NCBRF	Nisonger Child Behavior Rating Form
PS-A	social problem-solving then assertiveness
PSI	Parenting Stress Index
PTSD	post-traumatic stress disorder
RCT	randomised controlled trial
RR	risk ratio
SAS-ID	Zung Self-Rating Anxiety Scale for Adults with Intellectual Disabilities
SCL-90-R	Symptom CheckList-90-Revised
SF-12	12-Item Short Form Health Survey
SIB-R	Scales of Independent Behavior-Revised
SMD	standardised mean difference
SNAP-IV	Swanson, Nolan and Pelham Questionnaire - revised
TAU	treatment as usual

N.1 Psychological/psychosocial interventions

N.1.1 Psychological interventions versus control for mental health problems

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	psychological interventions	control	Relative (95% CI)	Absolute (95% CI)		
Mental health – RCTs (follow up: mean 13.25 weeks; assessed with: various scales)												
3	randomised trials	very serious [1]	serious [2]	not serious	serious [3]	none	41	-	-	SMD **1.24 SD lower** (2.31 lower to 0.18 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Mental health – Controlled before-and-after studies (follow up: 12 weeks; assessed with: Brief Symptom Inventory: Global Severity Index [GSI])												
1	observational studies	very serious [4]	not serious	not serious	serious [3]	none	12	12	-	MD **0.83 lower** (1.29 lower to 0.37 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Low problem behaviour (follow up: 10 weeks; assessed with: Role-play test of anger arousing situations)												
1	randomised trials	serious [1]	not serious	not serious	serious [5]	none	18	10	-	MD **11.69 more** (7.06 more to 16.32 more)	⨁⨁◯◯ LOW	IMPORTANT

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	psychological interventions	control	Relative (95% CI)	Absolute (95% CI)		
Maladaptive functioning (follow up: 10 weeks; assessed with: Adaptive Behaviour Scale – revised – part II)												
1	randomised trials	serious [1]	not serious	not serious	serious [3]	none	18	10	-	MD **21.74 lower** (36.45 lower to 7.02 lower)	⊕⊕◯◯ LOW	IMPORTANT
Interpersonal skills (follow up: 18 weeks; assessed with: Social Performance Survey Schedule)												
1	randomised trials	serious [1]	not serious	not serious	very serious [6]	none	22	10	-	MD **20.45 more** (9.74 fewer to 50.74 more)	⊕◯◯◯ VERY LOW	IMPORTANT

1. Risk of selection and performance bias
2. I2 suggests considerable heterogeneity
3. Confidence intervals cross minimally important difference in one direction. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes)
4. Risk of selection and performance bias and unclear risk of selective outcomes, attrition and detection bias
5. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
6. Confidence intervals cross two minimally important differences. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.1.2 Social problem-solving than assertiveness training (PS-A) versus assertiveness then social problem-solving (A-PS) for mental health problems

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	social problem-solving, then assertiveness training (PS-A)	assertiveness, then social problem-solving (A-PS)	Relative (95% CI)	Absolute (95% CI)		
Psychiatric/psychological symptoms (follow up: 23 weeks; assessed with: Brief Symptom Inventory)												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	social problem-solving, then assertiveness training (PS-A)	assertiveness, then social problem-solving (A-PS)	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious [1,2]	not serious	not serious	serious [3]	none	9	9	-	MD 0.02 more (0.43 fewer to 0.47 more)	⨁⨁◯◯ LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Psychological distress (follow up: 23 weeks; assessed with: Subjective Unit of Distress Scale)												
1	randomised trials	serious [1,2]	not serious	not serious	very serious [4]	none	9	9	-	MD 0.22 fewer (2.82 fewer to 2.38 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Low problem behaviour – Follow-up (follow up: 23 weeks; assessed with: Role-play test of anger arousing situations)												
1	randomised trials	serious [1,2]	not serious	not serious	serious [3]	none	9	9	-	MD 4.11 more (1.07 fewer to 9.29 more)	⨁⨁◯◯ LOW	IMPORTANT
Adaptive behaviour (follow up: 23 weeks; assessed with: Adaptive Behavior Scale – Revised)												
1	randomised trials	serious [1,2]	not serious	not serious	very serious [4]	none	9	9	-	MD 2.02 fewer (18.88 fewer to 14.84 more)	⨁◯◯◯ VERY LOW	IMPORTANT

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	social problem-solving, then assertiveness training (PS-A)	assertiveness, then social problem-solving (A-PS)	Relative (95% CI)	Absolute (95% CI)		
Adaptive behaviour (follow up: 23 weeks; assessed with: Problem-Solving Task)												
1	randomised trials	serious [1,2]	not serious	not serious	very serious [4]	none	9	9	-	MD 4 fewer (20.7 fewer to 12.7 more)	⨁◯◯◯ VERY LOW	IMPORTANT

1. Risk of selection bias (unclear allocation method, no details of allocation concealment)
2. Risk of performance bias (not blind)
3. Confidence intervals cross one minimally important difference. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
4. Confidence intervals cross minimally important difference in both directions (downgrade 2). Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.1.3 Psychodynamic psychotherapy (8 sessions) versus psychodynamic psychotherapy (12 or 24+ sessions) for mental health problems

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	psychodynamic psychotherapy (8 sessions)	psychodynamic psychotherapy (12 or 24+ sessions)	Relative (95% CI)	Absolute (95% CI)		
Mental health (follow up: ?; assessed with: SCL-90-R)												
1	observational studies	very serious [1]	not serious	not serious	serious [2]	none	No statistically significant differences were found between arms with differing lengths of treatment				⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	psychodynamic psychotherapy (8 sessions)	psychodynamic psychotherapy (12 or 24+ sessions)	Relative (95% CI)	Absolute (95% CI)		
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Interpersonal problems (follow up: ?; assessed with: Inventory of Interpersonal Problems-32)												
1	observational studies	very serious [1]	not serious	not serious	serious [2]	none	No statistically significant differences were found between arms with differing lengths of treatment				⨁◯◯◯ VERY LOW	IMPORTANT

1. Risk of selection, detection and performance bias.
2. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.1.4 Psychological interventions versus control for substance misuse

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	psychological interventions	control	Relative (95% CI)	Absolute (95% CI)		
Alcohol misuse (follow up: 34 weeks)												
1	randomised trials	very serious [1]	not serious	not serious	very serious [2]	none	42	42	-	MD **0.12 fewer** (1.01 fewer to 0.77 more)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	psychological interventions	control	Relative (95% CI)	Absolute (95% CI)		
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL

1. Risk of selection bias (no details of allocation method or concealment but, most importantly and not comparable risk at baseline), risk of performance bias
2. Confidence intervals cross minimally important difference in both directions (downgrade 2). Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.1.5 Assertiveness training versus modelling and social inference for substance misuse

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	assertiveness training	modelling and social inference	Relative (95% CI)	Absolute (95% CI)		
Alcohol misuse (follow up: mean 34 weeks)												
1	randomised trials	very serious [1]	not serious	not serious	very serious [2]	none	21	21	-	MD **0.07 fewer** (0.82 fewer to 0.68 more)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL

1. Risk of selection bias (no details of allocation method or concealment but, most importantly, not comparable risk at baseline), Risk of performance bias
2. Confidence intervals cross minimally important difference in both directions (downgrade 2). Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.1.6 Psychological intervention versus control for anxiety symptoms

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	psychological intervention	control	Relative (95% CI)	Absolute (95% CI)		
Anxiety symptoms (RCTs) (follow up: mean 42 weeks; assessed with: various scales)												
2	randomised trials	very serious [1]	serious [2]	not serious	very serious [3]	none	29	-	-	SMD **0.87 SD fewer** (1.14 fewer to 1.36 more)	⨁◯◯◯ VERY LOW	CRITICAL
Anxiety symptoms (Controlled before-and-after) (follow up: 12 weeks; assessed with: Brief Symptom Inventory: anxiety symptom dimension)												
1	before-after studies	very serious [4]	not serious	not serious	serious [5]	none	12	12	-	MD **0.4 SD lower** (1.23 lower to 0.43 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
In paid employment after treatment (follow up: 16 weeks)												
1	randomised trials	very serious [6]	not serious	not serious	serious [5]	none	1/16 (6.3%)	4/14 (28.6%)	RR 0.22 (0.03 to 1.73)	**223 fewer per 1000** (from 209 more to 277 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Voluntary work (follow up: 16 weeks)												
1	randomised trials	very serious [6]	not serious	not serious	very serious [3]	none	6/16 (37.5%)	4/14 (28.6%)	RR 1.31 (0.46 to 3.72)	**89 more per 1000** (from 154 fewer to 777 more)	⨁◯◯◯ VERY LOW	CRITICAL

1. Risk of selection, performance and detection bias
2. I2 suggests considerable heterogeneity
3. Confidence intervals cross minimally important difference in both directions. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes)
4. Risk of selection and performance bias and unclear risk of attrition and detection bias
5. Confidence intervals cross minimally important difference in one direction. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes)
6. Risk of performance and selection bias

N.1.7 Relaxation training versus control for anxiety symptoms

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	relaxation training	control	Relative (95% CI)	Absolute (95% CI)		
Anxiety symptoms (Group relaxation training versus control) (follow up: range 2.29 weeks to unclear; assessed with: various tools)												
2	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	35	-	-	SMD **2.31 lower** (2.92 lower to 1.7 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Anxiety symptoms (Individual relaxation training versus control) (follow up: 2.29 weeks; assessed with: 5-point scale on 10 ratings; Scale from: relaxed to very anxious)												
2	randomised trials	very serious [3]	serious [4]	not serious	serious [2]	none	20	-	-	SMD **2.97 SD lower** (4.36 lower to 1.57 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (relaxation versus story-telling) – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation (relaxation versus story-telling) – not reported												
-	-	-	-	-	-	-					-	CRITICAL

1. Risk of selection, performance and possible detection bias
2. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes)
3. Risk of selection bias (no details of allocation method or concealment); Risk of performance bias (no blinding); Possible risk of detection bias (unclear if outcome assessors blind to treatment and confounding)
4. I2 suggests substantial heterogeneity.

N.1.8 Dating skills versus control for social anxiety symptoms

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	dating skills training	control	Relative (95% CI)	Absolute (95% CI)		
Social anxiety symptoms (follow up: 24 weeks; assessed with: Social Avoidance and Distress Scale)												
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	13	12	-	MD **0.39 lower** (1.18 lower to 0.4 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Proportion with significant change in social anxiety symptoms (follow up: 20 weeks; assessed with: Social Avoidance and Distress Scale)												
1	randomised trials	very serious [1]	not serious	not serious	serious [3]	none	-/13	-/12	not estimable		⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL

1. Risk of selection and detection bias
2. Confidence intervals cross one minimally important difference. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
3. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.1.9 Cognitive behavioural therapy versus ABA/IBI for post-traumatic stress disorder

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	ABA/IBI	Relative (95% CI)	Absolute (95% CI)		
Somatic symptoms (follow up: not reported; assessed with: Achenbach: somatic subscale)												
1	before-after studies	very serious [1]	not serious	not serious	serious [2]	none	42	45	-	MD **3.74 more** (0.69 more to 6.79 more)	⨁◯◯◯ VERY LOW	CRITICAL
Withdrawn symptoms (follow up: not reported; assessed with: Achenbach: withdrawn subscale)												
1	before-after studies	very serious [1]	not serious	not serious	serious [2]	none	42	45	-	MD **4.58 more** (1.12 more to 8.04 more)	⨁◯◯◯ VERY LOW	CRITICAL
Anxious/depressed symptoms (follow up: not reported; assessed with: Achenbach: anxious/depressed subscale)												
1	before-after studies	very serious [1]	not serious	not serious	serious [2]	none	42	45	-	MD **6.89 more** (3.68 more to 10.1 more)	⨁◯◯◯ VERY LOW	CRITICAL
Thought problems (follow up: not reported; assessed with: Achenbach: thought problems subscale)												
1	before-after studies	very serious [1]	not serious	not serious	very serious [3]	none	42	45	-	MD **7.53 more** (4.83 more to 10.23 more)	⨁◯◯◯ VERY LOW	CRITICAL
Attention subscale (follow up: not reported; assessed with: Achenbach: attention subscale)												
1	before-after studies	very serious [1]	not serious	not serious	serious [2]	none	42	45	-	MD **4.58 more** (1.56 more to 7.6 more)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	ABA/IBI	Relative (95% CI)	Absolute (95% CI)		
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Social problems (follow up: not reported; assessed with: Achenbach: social problems subscale)												
1	before-after studies	very serious [1]	not serious	not serious	serious [2]	none	42	45	-	MD **2.97 more** (0.38 fewer to 6.32 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Aggressive behaviour (follow up: not reported; assessed with: Achenbach: Aggressive behaviour subscale)												
1	before-after studies	very serious [1]	not serious	not serious	very serious [3]	none	42	45	-	MD **7.22 more** (4.66 more to 9.78 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Rule breaking symptoms (follow up: not reported; assessed with: Achenbach: Rule breaking subscale)												
1	before-after studies	very serious [1]	not serious	not serious	very serious [3]	none	42	45	-	MD **9.18 more** (6.95 more to 11.41 more)	⨁◯◯◯ VERY LOW	IMPORTANT

1. Risk of selection bias, performance bias (no blinding) and unclear risk of attrition bias
2. Confidence intervals cross minimally important difference in one direction. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes)
3. Confidence intervals cross minimally important difference in both directions. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes)

N.1.10 Cognitive behavioural therapy versus control for depressive symptoms

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	control	Relative (95% CI)	Absolute (95% CI)		
Depressive symptoms (RCT) (follow up: range 6 weeks to 42 weeks; assessed with: BDI)												
3	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	68	-	-	SMD **0.82 fewer** (1.64 fewer to 0)	⨁◯◯◯ VERY LOW	CRITICAL
Depressive symptoms (Controlled before-and-after) (follow up: range 12 weeks to 46.7 weeks; assessed with: various)												
3	observational studies	very serious [3]	not serious	not serious	serious [2]	none	84	-	-	SMD **0.81 lower** (1.39 lower to 0.23 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Depression: at least small improvement (follow up: 12 weeks; assessed with: BDI)												
1	randomised trials	serious [4]	not serious	not serious	serious [2]	none	19/20 (95.0%)	17/27 (63.0%)	RR 1.51 (1.11 to 2.05)	**321 more per 1000** (from 69 more to 661 more)	⨁⨁◯◯ LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
In paid employment after treatment (follow up: 16 weeks)												
1	randomised trials	very serious [5]	not serious	not serious	serious [2]	none	-/16	4/14 (28.6%)	RR 0.22 (0.03 to 1.73)	**223 fewer per 1000** (from 209 more to 277 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
In voluntary work after treatment (follow up: 16 weeks)												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious [5]	not serious	not serious	very serious [6]	none	-/16	4/14 (28.6%)	RR 1.31 (0.46 to 3.72)	**89 more per 1000** (from 154 fewer to 777 more)	⨁◯◯◯ VERY LOW	CRITICAL
Problem behaviour (Controlled before-and-after) (follow up: 23 weeks; assessed with: SIB-R)												
1	before-after studies	very serious [3]	not serious	not serious	serious [2]	none	16	8	-	MD **7 fewer** (18.58 fewer to 4.58 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Social skills (mild to moderate learning disabilities) (follow up: 6-12 weeks; assessed with: Social comparison scale)												
2	randomised trials	very serious [5]	serious [7]	not serious [8]	serious [2]	none	54	42	-	MD **1.24 more** (0.66 more to 1.82 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Social behaviours (Controlled before-and-after) (follow up: 23 weeks; assessed with: Social performance survey schedule)												
1	before-after studies	very serious [3]	serious [8]	not serious	serious [9]	none	16	8	-	MD **11.12 fewer** (17.11 fewer to 5.13 fewer)	⨁◯◯◯ VERY LOW	IMPORTANT

; **RR:** Risk ratio; **MD:** Mean difference

1. Risk of selection and performance bias in studies contributing to >50% weighting in analysis
2. Confidence intervals cross one minimally important difference. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
3. Risk of selection, performance and detection bias
4. Risk of selection bias
5. Risk of selection and performance bias
6. Confidence intervals cross minimally important differences in both directions. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
7. No explanation was provided
8. Inconsistency in the impact on social skills between RCTs and controlled before-and-after studies.
9. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.1.11 Cognitive behavioural therapy versus behavioural strategies for depressive symptoms

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	behavioural strategies only	Relative (95% CI)	Absolute (95% CI)		
Depressive symptoms (follow up: 38 weeks; assessed with: BDI-II)												
1	observational studies	very serious [1]	not serious	not serious	serious [2]	none	23	24	-	MD **1.56 fewer** (6.57 fewer to 3.45 more)	⨁◯◯◯ VERY LOW	CRITICAL
Improvement in those with clinical depression at baseline (follow up: 38 weeks; assessed with: BDI-II [reduced score])												
1	observational studies	very serious [1]	not serious	not serious	serious [2]	none	14/14 (100.0%)	14/17 (82.4%)	RR 1.20 (0.94 to 1.53)	**165 more per 1000** (from 49 fewer to 436 more)	⨁◯◯◯ VERY LOW	CRITICAL
Recovery in those with clinical depression at baseline (follow up: 38 weeks; assessed with: BDI-II [score 12 or less])												
1	observational studies	very serious [1]	not serious	not serious	very serious [3]	none	8/14 (57.1%)	12/17 (70.6%)	RR 0.81 (0.47 to 1.40)	**134 fewer per 1000** (from 282 more to 374 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL

1. Risk of selection, performance and detection bias
2. Confidence intervals cross minimally important difference in one direction. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes)
3. Confidence intervals cross minimally important difference in both directions. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes)

N.1.12 Cognitive behavioural therapy versus cognitive strategies for depressive symptoms

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CBT	cognitive strategies only	Relative (95% CI)	Absolute (95% CI)		
Depressive symptoms (follow up: 38 weeks; assessed with: BDI-II)												
1	observational studies	very serious [1]	not serious	not serious	serious [2]	none	23	23	-	MD **1.3 fewer** (5.89 fewer to 3.29 more)	⨁◯◯◯ VERY LOW	CRITICAL
Improvement in those with clinical depression at baseline (follow up: 38 weeks; assessed with: BDI-II [reduced score])												
1	observational studies	very serious [1]	not serious	not serious	serious [2]	none	14/14 (100.0%)	11/15 (73.3%)	RR 1.34 (0.98 to 1.85)	**249 more per 1000** (from 15 fewer to 623 more)	⨁◯◯◯ VERY LOW	CRITICAL
Recovery in those with clinical depression at baseline (follow up: 38 weeks; assessed with: BDI-II [score 13 or less])												
1	observational studies	very serious [1]	not serious	not serious	very serious [3]	none	8/14 (57.1%)	7/15 (46.7%)	RR 1.22 (0.60 to 2.48)	**103 more per 1000** (from 187 fewer to 691 more)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL

1. Risk of selection, performance and detection bias
2. Confidence intervals cross minimally important difference in one direction. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes)
3. Confidence intervals cross minimally important difference in both directions. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes)

N.1.13 Psychodynamic psychotherapy versus no treatment for sexually inappropriate behaviour

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	psychodynamic psychotherapy	no treatment	Relative (95% CI)	Absolute (95% CI)		
Recidivism (follow up: 208 weeks)												
1	observational studies	serious [1]	not serious	serious [2]	very serious [3]	none	2/13 (15.4%)	3/5 (60.0%)	**RR 0.26** (0.06 to 1.11)	**444 fewer per 1000** (from 66 more to 564 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL

1. Risk of selection bias, performance bias
2. Participants are only those who were arrested by the criminal justice system and, therefore, are unlikely to represent all individuals with learning disabilities who present with sexually inappropriate behaviour as not all will be in contact with the criminal justice system.
3. Confidence intervals cross minimally important difference in both directions. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes)

N.1.14 Parent training versus control

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	parent training	any control	Relative (95% CI)	Absolute (95% CI)		
Behavioural and emotional problems (severity) – post-treatment (assessed with: various scales)												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**parent training**	**any control**	**Relative (95% CI)**	**Absolute (95% CI)**		
13	randomised trials	serious [1]	not serious	not serious	not serious	none	349	-	-	SMD **0.4 SD lower** (0.55 lower to 0.24 lower)	⊕⊕⊕◯ MODERATE	CRITICAL
Behavioural and emotional problems (severity) – follow-up (follow up: range 26- 52 weeks to 0; assessed with: various scales)												
2	randomised trials	serious [1]	not serious	serious [2]	serious [3]	publication bias strongly suspected	86	-	-	SMD **0.13 fewer** (0.45 fewer to 0.19 more)	⊕◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Problem behaviour (severity, non-improvement) – post-treatment (assessed with: various scales)												
8	randomised trials	serious [1]	not serious	not serious	not serious	none	131/231 (56.7%)	174/197 (88.3%)	RR 0.67 (0.59 to 0.77)	**291 fewer per 1000** (from 203 fewer to 362 fewer)	⊕⊕⊕◯ MODERATE	IMPORTANT
Problem behaviour (frequency) – post-treatment (assessed with: various scales)												
8	randomised trials	serious [1]	serious [4]	not serious	not serious	none	237	-	-	SMD **0.6 fewer** (0.9 fewer to 0.3 fewer)	⊕⊕◯◯ LOW	IMPORTANT
Problem behaviour (frequency) – follow-up (follow up: mean 26 weeks; assessed with: various scales)												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	parent training	any control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious [5]	not serious	not serious	very serious [6]	publication bias strongly suspected	35	-	-	SMD **0.36 fewer** (0.85 fewer to 0.14 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Problem behaviour (frequency, non-improvement) – post-treatment (assessed with: various scales)												
6	randomised trials	serious [1]	not serious	serious [2]	not serious	none	105/188 (55.9%)	147/155 (94.8%)	RR 0.63 (0.55 to 0.73)	**351 fewer per 1000** (from 256 fewer to 427 fewer)	⨁⨁◯◯ LOW	IMPORTANT
Adaptive functioning (communication) – post-treatment												
1	randomised trials	serious [5]	not serious	serious [2]	very serious [6]	none	75	-	-	SMD **0.47 more** (0.11 more to 0.84 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Adaptive functioning (total) – post-treatment												
2	randomised trials	serious [1]	not serious	serious [2]	serious [3]	none	82	-	-	SMD **0.51 more** (0.15 more to 0.86 more)	⨁◯◯◯ VERY LOW	IMPORTANT

1. Most information is from studies at moderate risk of bias
2. Concerns with applicability – different populations
3. Optimal information size not met
4. I2 > 40%. This is the criterion that was used in the challenging behaviour guideline.
5. Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
6. Optimal information size not met; small, single study
7. Publication bias strongly suspected

For the full GRADE evidence profiles for other pairwise comparisons relating to the quality of evidence for parent training, please refer to the NICE guideline Challenging Behaviour and Learning Disabilities, NG11.

N.2 Pharmacological interventions

N.2.1 Amphetamine versus placebo

Quality assessment							Impact	Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Overall effect of treatment on bespoke form (follow up: mean 23 weeks; assessed with: 14-item 'patient evaluation form')									
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	The differences between groups on 10 subscales (hyperkinesis, concentration, attention, aggressiveness, sociability, interpersonal relationship, mood, work capacity, reading, spelling, arithmetic and class standing) were reported as not significant; however, the comprehension and work interest subscales were reported to be significantly better in the amphetamine group than the placebo group ($p < 0.05$).	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported									
-	-	-	-	-	-	-		-	CRITICAL
Community participation and meaningful occupation – not reported									
-	-	-	-	-	-	-		-	CRITICAL

1. Risk of selection and selective outcomes bias; unclear risk of detection, attrition and performance bias.
2. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.2.2 Methylphenidate versus placebo

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Methylphenidate	placebo	Relative (95% CI)	Absolute (95% CI)		
ADHD (follow up: mean 16 weeks; assessed with: Connors' ADHD index [parent rated])												
1	randomised trials	not serious	not serious	not serious	serious [1]	none	61	61	-	MD **3.3 fewer** (6.79 fewer to 0.19 more)	⊕⊕⊕◯ MODERATE	CRITICAL
ADHD (follow up: mean 16 weeks; assessed with: Connors' ADHD index [teacher rated])												
1	randomised trials	not serious	not serious	not serious	serious [1]	none	61	61	-	MD **4.1 fewer** (7.57 fewer to 0.63 fewer)	⊕⊕⊕◯ MODERATE	CRITICAL
Hyperactivity (follow up: mean 16 weeks; assessed with: Conners' hyperactivity scale [parent rated])												
1	randomised trials	not serious	not serious	not serious	serious [2]	none	61	61	-	MD **1.5 fewer** (3.44 fewer to 0.44 more)	⊕⊕⊕◯ MODERATE	CRITICAL
Hyperactivity (follow up: mean 16 weeks; assessed with: Conners' hyperactivity scale [teacher rated])												
1	randomised trials	not serious	not serious	not serious	serious [1]	none	61	61	-	MD **2.6 fewer** (4.68 fewer to 0.52 fewer)	⊕⊕⊕◯ MODERATE	CRITICAL
'Improved' or 'better' (follow up: mean 16 weeks; assessed with: Clinical Global Impressions-Improvement)												
1	randomised trials	not serious	not serious	not serious	serious [1]	none	24/61 (39.3%)	4/61 (6.6%)	RR 6.00 (2.21 to 16.26)	**328 more per 1000** (from 79 more to 1000 more)	⊕⊕⊕◯ MODERATE	CRITICAL
Quality of life – not reported												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Methylphenidate	placebo	Relative (95% CI)	Absolute (95% CI)		
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Weight (follow up: mean 16 weeks; assessed with: kg)												
1	randomised trials	not serious	not serious	not serious	serious [2]	none	61	61	-	MD **4.2 kg fewer** (10.25 fewer to 1.85 more)	⊕⊕⊕◯ MODERATE	IMPORTANT

1. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
2. Confidence intervals cross one minimally important difference. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.2.3 Methylphenidate plus behavioural modification training vs placebo plus behavioural modification training

Quality assessment							Impact	Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Behaviour (including ADHD and hyperactivity) (follow up: 2 weeks; assessed with: Conner's Teacher Report form - all subscales)									
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	The authors found significant improvement for methylphenidate treatment compared to placebo on two categories: behaviour modification and deviant vocalization. However, they reported that this only occurred when the behavioural modification program was in place.	⊕◯◯◯ VERY LOW	CRITICAL
Quality of life - not reported									
-	-	-	-	-	-	-		-	

Quality assessment							Impact	Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Community participation and meaningful occupation - not reported									
-	-	-	-	-	-	-		-	

CI: Confidence interval

1. Risk of selection and detection bias.
2. Sample size less than optimal information size (<400 for continuous outcomes or <300 events for dichotomous outcomes).

N.2.4 Clonidine versus placebo

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Clonidine	placebo	Relative (95% CI)	Absolute (95% CI)		
ADHD symptoms: conduct (follow up: 6 weeks; assessed with: Parent Connor's score – conduct scale)												
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	9	10	-	MD **7.4 fewer** (10.34 fewer to 4.46 fewer)	⨁◯◯◯ VERY LOW	
ADHD symptoms: impulsive hyperactivity (follow up: 6 weeks; assessed with: Parent Connor's score – Impulsive hyperactive scale)												
1	randomised trials	very serious [1]	not serious	not serious	serious [3]	none	9	10	-	MD **2.6 fewer** (6.54 fewer to 1.34 more)	⨁◯◯◯ VERY LOW	
ADHD symptoms: overall (follow up: 6 weeks; assessed with: Parent Connor's score – Total score)												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Clonidine**	**placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
1	randomised trials	very serious [1]	not serious	not serious	serious [3]	none	9	10	-	MD **24.7 fewer** (49.35 fewer to 0.05 fewer)	⨁◯◯◯ VERY LOW	
ADHD symptoms (clinician rated) (follow up: 6 weeks; assessed with: CGI)												
1	randomised trials	very serious [1]	not serious	not serious	serious [3]	none	9	10	-	MD **1.8 fewer** (3.11 fewer to 0.49 fewer)	⨁◯◯◯ VERY LOW	
Much or very much improved (follow up: 6 weeks; assessed with: CGI)												
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	7/9 (77.8%)	0/10 (0.0%)	RR 16.50 (1.07 to 253.40)	**0 fewer per 1000**[4] (from 0 fewer to 0 fewer)	⨁◯◯◯ VERY LOW	
Quality of life – not reported												
-	-	-	-	-	-	-					-	
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	

1. Risk of selection and selective outcome reporting bias
2. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
3. Confidence intervals cross one minimally important difference. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
4. Absolute risk value is 0 as no events of interest occurred for this outcome

N.2.5 Risperidone versus methylphenidate

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Risperidone	methylphenidate	Relative (95% CI)	Absolute (95% CI)		
ADHD symptoms (follow up: mean 4 weeks; assessed with: SNAP-IV total score)												
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	22	-	-	SMD **0.54 lower** (1.14 lower to 0.06 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Hyperactivity (NCBRF) (follow up: mean 4 weeks)												
1	randomised trials	very serious [1]	not serious	not serious	serious [3]	none	No significant between-group differences in change scores.				⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Side effects (Barkley's Side Effects Rating Scale) (follow up: mean 4 weeks)												
1	randomised trials	very serious [1]	not serious	not serious	very serious [2]	none	22	-	-	SMD **0.08 higher** (0.54 lower to 0.69 higher)	⨁◯◯◯ VERY LOW	IMPORTANT
Weight (follow up: 4 weeks; assessed with: kg)												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Risperidone	methylphenidate	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious [1]	not serious	not serious	serious [3]	none	Mean reduction of 0.53 kg in the methylphenidate group compared with a weight increase of 1.01 kg in the risperidone group (reported to be significant).				⨁◯◯◯ VERY LOW	

1. Risk of selection and selective outcome reporting bias
2. Confidence intervals cross one minimally important difference. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
3. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.3 Pharmacological interventions for dementia in Down's syndrome

N.3.1 Donepezil versus placebo for prevention of dementia

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Comparison 1a: donepezil	placebo	Relative (95% CI)	Absolute (95% CI)		
Cognitive abilities (follow up: 12 weeks; assessed with: Severe Impairment Battery)												
2	randomised trials	not serious	very serious [1]	not serious [2]	very serious [1]	none	68	-	-	SMD **0.34 higher** (0.65 lower to 1.33 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Comparison 1a: donepezil	placebo	Relative (95% CI)	Absolute (95% CI)		
-	-	-	-	-	-	-					-	CRITICAL
Behavioural problems (follow up: 12 weeks; assessed with: various scales)												
2	randomised trials	not serious	not serious	not serious	serious [3]	none	62	-	-	SMD **0.28 higher** (0.07 lower to 0.63 higher)	⊕⊕⊕◯ MODERATE	IMPORTANT
Serious adverse events (follow up: 12 weeks)												
2	randomised trials	not serious	not serious	not serious	serious [4]	none	0/71 (0.0%)	0/70 (0.0%)	not estimable		⊕⊕⊕◯ MODERATE	IMPORTANT
Severe adverse events (follow up: 12 weeks)												
1	randomised trials [5]	not serious	not serious	not serious	very serious [2]	none	2/62 (3.2%)	0/61 (0.0%)	RR 4.92 (0.24 to 100.43)	**0 fewer per 1000**[6] (from 0 fewer to 0 fewer)	⊕⊕◯◯ LOW	IMPORTANT
Any adverse event (follow up: 12 weeks)												
1	randomised trials [7]	not serious	not serious	not serious	serious [3]	none	46/62 (74.2%)	29/61 (47.5%)	RR 1.56 (1.15 to 2.11)	**266 more per 1000** (from 71 more to 528 more)	⊕⊕⊕◯ MODERATE	IMPORTANT

1. Downgraded two levels for imprecision (wide confidence interval) and inconsistency (I^2 = 73%). This was the criterion used in the Livingstone 2015 review.
2. Downgraded two levels for serious imprecision (wide confidence interval) and small number of events. This was the criterion used in the Livingstone 2015 review.
3. Downgraded one level for imprecision (wide confidence interval). This was the criterion used in the Livingstone 2015 review.
4. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
5. Serious adverse events: hypertension and emotional lability.
6. Absolute risk value is 0 as no events of interest occurred for this outcome.
7. Most common side effects were asthenia, anorexia, dyspepsia, nausea, vomiting, and insomnia.

N.3.2 Donepezil versus placebo for treatment of dementia

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Comparison 1b: donepezil	placebo	Relative (95% CI)	Absolute (95% CI)		
Cognitive abilities (follow up: 24 weeks; assessed with: Severe Impairment Battery)												
1	randomised trials	not serious	not serious	not serious	serious [1]	none	14	-	-	SMD **0.93 higher** (0.13 higher to 1.73 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Proportion with improved impression of quality of life (follow up: 24 weeks)												
1	randomised trials	not serious	not serious	not serious	serious [1]	none	11/11 (100.0%)	4/10 (40.0%)	RR 2.34 (1.14 to 4.81)	**536 more per 1000** (from 56 more to 1000 more)	⊕⊕⊕◯ MODERATE	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Behavioural problems (follow up: 24 weeks; assessed with: American Association of Mental Retardation Adaptive Behaviour Scale)												
1	randomised trials	not serious	not serious	not serious	serious [1]	none	14	-	-	SMD **0.99 higher** (0.18 higher to 1.79 higher)	⊕⊕⊕◯ MODERATE	IMPORTANT
Serious adverse events (follow up: 24 weeks)												
1	randomised trials	not serious	not serious	not serious	serious [1]	none	8/16 (50.0%)	3/14 (21.4%)	RR 2.33 (0.76 to 7.13)	**285 more per 1000** (from 51 fewer to 1000 more)	⊕⊕⊕◯ MODERATE	IMPORTANT
At least one serious event (follow up: 24 weeks)												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Comparison 1b: donepezil	placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	serious [1]	none	12/16 (75.0%)	7/14 (50.0%)	RR 1.50 (0.83 to 2.72)	**250 more per 1000** (from 85 fewer to 860 more)	⨁⨁⨁◯ MODERATE	IMPORTANT
Minor adverse reaction (follow up: 24 weeks)												
1	randomised trials [2]	not serious	not serious	not serious	very serious [3]	none	2/11 (18.2%)	3/10 (30.0%)	RR 0.61 (0.13 to 2.92)	**117 fewer per 1000** (from 261 fewer to 576 more)	⨁⨁◯◯ LOW	IMPORTANT

1. Downgraded one level for imprecision (wide confidence interval). This was the criterion used in the Livingstone 2015 review.
2. Included soft stool and skin rash (donepezil, one placebo) or mild skin rash only (2 placebo).
3. Downgraded two levels for serious imprecision (wide confidence interval).

N.3.3 Memantine versus placebo for dementia in Down's syndrome

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Comparison 2: memantine	placebo	Relative (95% CI)	Absolute (95% CI)		
Cognitive abilities (follow up: range 16 weeks to 52 weeks; assessed with: various scales)												
2	randomised trials	not serious	serious [1]	not serious	serious [1]	none	91	-	-	SMD **0.05 more** (0.43 fewer to 0.52 more)	⨁⨁◯◯ LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Comparison 2: memantine	placebo	Relative (95% CI)	Absolute (95% CI)		
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Behavioural problems (follow up: range 16 weeks to 52 weeks; assessed with: various scales)												
2	randomised trials	not serious	not serious	not serious	very serious [2]	none	94	-	-	SMD **0.17 fewer** (0.46 fewer to 0.11 more)	⊕⊕◯◯ LOW	IMPORTANT
Clinically significant/serious adverse events (follow up: range 16 weeks to 52 weeks)												
2	randomised trials	not serious	not serious	not serious	very serious [2]	none	12/107 (11.2%)	6/104 (5.8%)	RR 1.79 (0.72 to 4.50)	**46 more per 1000** (from 16 fewer to 202 more)	⊕⊕◯◯ LOW	IMPORTANT
Any adverse event (follow up: mean 16 weeks)												
1	randomised trials	not serious	not serious	not serious	very serious [2]	none	4/19 (21.1%)	1/19 (5.3%)	RR 4.00 (0.49 to 32.57)	**158 more per 1000** (from 27 fewer to 1000 more)	⊕⊕◯◯ LOW	IMPORTANT

1. Downgraded two levels due to imprecision (wide confidence intervals) and inconsistency (I^2 = 48%). This was the criterion used in the Livingstone 2015 review.
2. Downgraded two levels for serious imprecision (wide confidence interval) and small number of events. This was the criterion used in the Livingstone 2015 review.

N.3.4 Simvastatin versus placebo for dementia in Down's syndrome

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Comparison 3: simvastatin	placebo	Relative (95% CI)	Absolute (95% CI)		
Cognitive abilities (follow up: 52 weeks; assessed with: Neuropsychological Assessment of Dementia in Intellectual Disabilities battery)												
1	randomised trials	not serious	not serious	not serious	very serious [1]	none	10	11	-	MD **10 higher** (0.4 lower to 1.6 higher)	⊕⊕◯◯ LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Adaptive functioning (follow up: 52 weeks; assessed with: AAMR: ABS)												
1	randomised trials	not serious	not serious	not serious	very serious [1]	none	10	11	-	MD **0.7 higher** (0 to 1.4 higher)	⊕⊕◯◯ LOW	IMPORTANT

1. Downgraded two levels for serious imprecision (wide confidence interval) and small number of events.

N.4 Other interventions

N.4.1 Annual health check versus treatment as usual

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Annual health check	treatment as usual	Relative (95% CI)	Absolute (95% CI)		
Psychosis (Identification of mental health needs; all levels of learning disabilities) (follow up: mean 39 weeks)												
1	randomised trials	serious [1]	not serious	serious [2]	serious [3]	none	4/83 (4.8%)	6/66 (9.1%)	**RR 0.53** (0.16 to 1.80)	**43 fewer per 1000** (from 73 more to 76 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Psychiatric consultation/ visit (Identification of mental health needs; all levels of learning disabilities) (follow up: range 39 weeks to 52 weeks)												
2	randomised trials	serious [4]	not serious	serious [2]	very serious [5]	none	26/287 (9.1%)	31/287 (10.8%)	**RR 0.83** (0.50 to 1.36)	**18 fewer per 1000** (from 39 more to 54 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Psychiatric disorders (Identification of mental health needs; all levels of learning disabilities) (follow up: mean 52 weeks)												
1	randomised trials	serious [1]	not serious	serious [2]	very serious [5]	none	2/234 (0.9%)	0/219 (0.0%)	**RR 4.68** (0.23 to 96.96)	**0 fewer per 1000**[8] (from 0 fewer to 0 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Newly detected health issues (all levels of learning disabilities) (follow up: range 39 weeks to 52 weeks)												
3	randomised trials	serious [1]	not serious	serious [2]	serious [3]	none	-/367	-/352	**OR 1.69** (1.08 to 2.64)	**0 fewer per 1000**[9] (from 0 fewer to 0 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Newly detected health monitoring needs (all levels of learning disabilities) (follow up: mean 39 weeks)												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Annual health check	treatment as usual	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious [1]	not serious	serious [2]	serious [6]	none	-/83	-/66	**OR 2.38** (1.31 to 4.32)	**0 fewer per 1000** (from 0 fewer to 0 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Newly detected health promotion needs (all levels of learning disabilities) (follow up: mean 39 weeks)												
1	randomised trials	serious [1]	not serious	serious [2]	very serious [5]	none	-/83	-/66	**OR 0.98** (0.73 to 1.32)	**0 fewer per 1000** (from 0 fewer to 0 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Obesity (Identification of health needs; all levels of learning disabilities) (follow up: range 39 weeks to 52 weeks)												
2	randomised trials	serious [1]	serious [7]	serious [2]	serious [6]	none	74/317 (23.3%)	43/285 (15.1%)	**RR 1.41** (1.09 to 1.82)	**62 more per 1000** (from 14 more to 124 more)	⨁◯◯◯ VERY LOW	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL

1. Risk of performance bias
2. Indirect outcome
3. Confidence intervals cross one minimally important difference. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
4. Risk of performance, selection, attrition bias
5. Confidence intervals cross two minimally important differences. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
6. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
7. I2 suggests considerable heterogeneity
8. Absolute risk value is 0 as no events of interest occurred for this outcome.
9. Absolute risk value is listed as 0 as data were not reported by the authors.

N.4.2 Acetyl-L-carnitine versus placebo for attention deficit hyperactivity disorder

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acetyl-L-carnitine	placebo	Relative (95% CI)	Absolute (95% CI)		
ADHD (follow up: mean 52 weeks; assessed with: Conners' Parents)												
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	24	27	-	MD **2.8 fewer** (7.58 fewer to 1.98 more)	⨁◯◯◯ VERY LOW	CRITICAL
ADHD (follow up: mean 52 weeks; assessed with: Conners' Teachers)												
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	24	27	-	MD **0.5 more** (5.08 fewer to 6.08 more)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Adaptive functioning (post-treatment) (follow up: mean 52 weeks; assessed with: VABS – full scale)												
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	24	27	-	MD **8.2 more** (0.04 fewer to 16.44 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Adaptive functioning (follow up: mean 52 weeks; assessed with: VABS – socialization scale)												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acetyl-L-carnitine	placebo	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	24	27	-	MD **11.3 more** (2.18 more to 20.42 more)	⨁◯◯◯ VERY LOW	IMPORTANT

1. Risk of selection and detection bias
2. Confidence intervals cross one minimally important difference. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.4.3 Acetyl-L-carnitine versus placebo for dementia

Quality assessment							Impact	Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Cognitive functioning (mild to moderate learning disabilities) (follow up: mean 39 weeks; assessed with: Multiple measures)									
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	No significant difference between Acetyl-L-Carnitine and placebo groups for all measures.	⨁◯◯◯ VERY LOW	CRITICAL
Dementia: (mild to moderate learning disabilities) (follow up: mean 39 weeks; assessed with: Emotional disorder rating scale)									
1	randomised trials	very serious [3]	not serious	not serious	serious [2]	none	No significant difference between Acetyl-L-Carnitine and placebo groups	⨁◯◯◯ VERY LOW	CRITICAL
Dementia (mild to moderate learning disabilities) (follow up: mean 39 weeks; assessed with: Child behaviour checklist)									
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	No significant difference between Acetyl-L-Carnitine and placebo groups	⨁◯◯◯ VERY LOW	CRITICAL

Quality assessment							Impact	Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Quality of life – not reported									
-	-	-	-	-	-	-		-	CRITICAL
Community participation and meaningful occupation – not reported									
-	-	-	-	-	-	-		-	CRITICAL

1. Risk of selection, selective outcomes and attrition bias.
2. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
3. Risk of selection, selective outcomes, detection bias and attrition bias.

N.4.4 Antioxidant plus acetylcholinesterase inhibitor versus placebo plus acetylcholinesterase inhibitor for dementia

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Antioxidant	placebo	Relative (95% CI)	Absolute (95% CI)		
Mental health (all levels of learning disabilities) (follow up: mean 104 weeks; assessed with: DMR [sum of cognitive scores])												
1	randomised trials	serious [1]	not serious	not serious	serious [2]	none	No significant differences in DMR cognitive scores scores between antioxidant and placebo groups				⊕⊕◯◯ LOW	CRITICAL
Mental health (all levels of learning disabilities) (follow up: mean 104 weeks; assessed with: Severe impairment battery)												
1	randomised trials	serious [1]	not serious	not serious	serious [2]	none	No significant differences in Severe Impairment Battery scores between antioxidant and placebo groups				⊕⊕◯◯ LOW	CRITICAL

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Antioxidant	placebo	Relative (95% CI)	Absolute (95% CI)		
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Adaptive functioning (all levels of learning disabilities) (follow up: mean 104 weeks; assessed with: Brief Praxis Test)												
1	randomised trials	serious [1]	not serious	not serious	serious [2]	none	No significant differences in Brief Praxis Test scores between antioxidant and placebo groups				⨁⨁◯◯ LOW	IMPORTANT
Adaptive functioning (all levels of learning disabilities) (follow up: mean 104 weeks; assessed with: DMR [sum of social skills])												
1	randomised trials	serious [1]	not serious	not serious	serious [2]	none	No significant differences in DMR sum of social scores scores between antioxidant and placebo groups				⨁⨁◯◯ LOW	IMPORTANT
Adaptive functioning (all levels of learning disabilities) (follow up: mean 104 weeks; assessed with: Bristol Activities of Daily Living Scale)												
1	randomised trials	serious [1]	not serious	not serious	serious [2]	none	No significant differences in Bristol Activities of Daily Living Scale scores between antioxidant and placebo groups				⨁⨁◯◯ LOW	IMPORTANT
Any serious adverse event (incapacitation and/or inability to sustain daily activities) (all levels of learning disabilities) (follow up: mean 104 weeks; assessed with: [ITT/analysed as randomised])												
1	randomised trials	serious [1]	not serious	not serious	serious [2]	none	14/29 (48.3%)	11/29 (37.9%)	RR 1.27 (0.70 to 2.32)	**102 more per 1000** (from 114 fewer to 501 more)	⨁⨁◯◯ LOW	IMPORTANT

1. Risk of selective outcomes bias.
2. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.4.5 Exercise versus any control for anxiety symptoms

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Exercise	any control	Relative (95% CI)	Absolute (95% CI)		
Anxiety (mild learning disabilities) (follow up: mean 39 weeks; assessed with: Hamilton Anxiety Scale)												
1	randomised trials	very serious [1,2]	not serious	not serious	serious [3]	none	Significant decrease in total HAM-A scores in the aerobic and leisure groups only (no significant decrease was found for the vocational activities control group.)				⨁◯◯◯ VERY LOW	CRITICAL
Anxiety (mild to moderate learning disabilities) (follow up: mean 12 weeks; assessed with: Zung Self-rating anxiety scale (adapted for learning disabilities and named Self-rated Anxiety Scale or SAS-ID)												
1	randomised trials	very serious [1]	not serious	not serious	serious [4]	none	14	13	-	MD 6.62 fewer (7.97 fewer to 5.27 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL

1. Risk of selection, performance and detection bias
2. Risk of selective outcome (no variance reported so not possible to use in meta-analysis), performance and selection bias
3. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes). Not possible to assess confidence without variance.
4. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.4.6 Exercise versus painting control for depressive symptoms

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Exercise	painting control	Relative (95% CI)	Absolute (95% CI)		
Depressive symptoms (mild to moderate learning disabilities) (follow up: mean 12 weeks; assessed with: Zung Self-rating Depression Scale)												
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	14	13	-	MD **6.06 fewer** (7.25 fewer to 4.87 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL

1. Risk of selection, performance and detection bias
2. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.4.7 Exercise and education versus control for depressive symptoms

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Exercise and education	control	Relative (95% CI)	Absolute (95% CI)		
Depressive symptoms (mild to moderate learning disabilities) (follow up: mean 12 weeks; assessed with: Child Depression Inventory)												
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	32	21	-	MD **1.53 fewer** (3.29 fewer to 0.23 more)	⨁◯◯◯ VERY LOW	CRITICAL

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Exercise and education	control	Relative (95% CI)	Absolute (95% CI)		
Community participation and meaningful occupation (mild to moderate learning disabilities) (follow up: mean 12 weeks; assessed with: Communication integration scale)												
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	32	21	-	MD **0.78 fewer** (2.06 fewer to 0.5 more)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (mild-moderate learning disabilities) (follow up: mean 12 weeks; assessed with: Life Satisfaction Scale)												
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	32	21	-	MD **2.52 more** (0.87 fewer to 5.91 more)	⨁◯◯◯ VERY LOW	CRITICAL

1. Selection and detection bias
2. Confidence intervals cross one minimally important difference. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.5 Organisation and Service Delivery

N.5.1 Assertive community treatment versus standard community treatment

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Assertive community treatment	standard community treatment	Relative (95% CI)	Absolute (95% CI)		
Mental health (service user) - not reported												
-	-	-	-	-	-	-					-	CRITICAL

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Assertive community treatment	standard community treatment	Relative (95% CI)	Absolute (95% CI)		
Healthcare practitioner health and well-being - not reported												
-	-	-	-	-	-	-					-	CRITICAL
Quality of life (follow up: range 13 weeks to 26 weeks)												
2	randomised trials	serious [1]	not serious	not serious	serious [2]	none	25	25	-	SMD **0.2 lower** (0.75 lower to 0.36 higher)	⨁⨁◯◯ LOW	CRITICAL
Community participation and meaningful occupation - not reported												
-	-	-	-	-	-	-					-	CRITICAL
Problem behaviours - not reported												
-	-	-	-	-	-	-					-	CRITICAL
Global assessment of function (symptomatology) (follow up: range 13 weeks to 26 weeks)												
2	randomised trials	serious [1]	not serious	not serious	serious [2]	none	25	25	-	MD **0.76 lower** (6.07 lower to 4.55 higher)	⨁⨁◯◯ LOW	IMPORTANT
Global assessment of function (Disability) (follow up: range 13 weeks to 26 weeks)												

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Assertive community treatment	standard community treatment	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	serious [1]	not serious	not serious	serious [2]	none	25	25	-	MD **1.05 higher** (4.05 lower to 6.16 higher)	⊕⊕◯◯ LOW	IMPORTANT
Carer uplift/burden (follow up: range 13 weeks to 26 weeks)												
2	randomised trials	serious [1]	not serious	not serious	very serious [3]	none	25	25	-	MD **0.03 higher** (3.48 lower to 3.54 higher)	⊕◯◯◯ VERY LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference

1. Risk of performance bias.
2. Confidence intervals cross one minimally important difference. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
3. Confidence intervals cross two minimally important differences. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.5.2 Active case management model versus standard model

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Active treatment case management model	standard model of service delivery	Relative (95% CI)	Absolute (95% CI)		
Mental health (service user) – not reported												
-	-	-	-	-	-	-				-	-	CRITICAL
Healthcare practitioner health and well-being – not reported												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Active treatment case management model	standard model of service delivery	Relative (95% CI)	Absolute (95% CI)		
-	-	-	-	-	-	-					-	CRITICAL
Quality of life (service user) – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Maladaptive behaviour (follow up: 3 years; assessed with: AAMD Maladaptive Behaviour Scale)												
1	randomised trials	very serious [1]	not serious	serious [2]	serious [3]	none	23	23	-	MD **12.91 fewer** (27.37 fewer to 1.55 more)	⨁◯◯◯ VERY LOW	CRITICAL
Adaptive behaviour (follow up: 3 years; assessed with: AAMD Adaptive Behaviour Scale)												
1	randomised trials	very serious [1]	not serious	serious [2]	serious [3]	none	23	23	-	MD **10.56 more** (6.77 fewer to 27.89 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Move to more staff intensive residential programming (follow up: 3 years)												
1	randomised trials	very serious [1]	not serious	serious [2]	very serious [4]	none	1/23 (4.3%)	4/23 (17.4%)	RR 0.25 (0.03 to 2.07)	**130 fewer per 1000** (from 169 fewer to 186 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Move to more staff intensive day programming (follow up: 3 weeks)												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Active treatment case management model	standard model of service delivery	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious [1]	not serious	serious [2]	very serious [4]	none	0/23 (0.0%)	2/23 (8.7%)	RR 0.20 (0.01 to 3.95)	**70 fewer per 1000** (from 86 fewer to 257 more)	⨁◯◯◯ VERY LOW	IMPORTANT

1. Risk of selection, performance and detection bias
2. American study so service structures less applicable to UK population
3. Confidence intervals cross one minimally important difference. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
4. Confidence intervals cross two minimally important differences. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.5.3 Liaison worker versus no liaison worker

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Liaison worker model	no liaison worker	Relative (95% CI)	Absolute (95% CI)		
Mental health (follow up: 39 weeks; assessed with: Strength and Difficulties Questionnaire)												
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	14	-	-	SMD **1.12 SD lower** (1.95 lower to 0.29 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (service user) – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Community participation and meaningful occupation – not reported												
-	-	-	-	-	-	-					-	CRITICAL

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Liaison worker model	no liaison worker	Relative (95% CI)	Absolute (95% CI)		
Problem behaviours – not reported												
-	-	-	-	-	-	-					-	CRITICAL
Carer quality of life – physical (follow up: 39 weeks; assessed with: SF-12-physical)												
1	randomised trials	very serious [1]	not serious	not serious	serious [2]	none	14	-	-	SMD **0.8 lower** (1.6 lower to 0)	⨁◯◯◯ VERY LOW	IMPORTANT
Care quality of life – mental (follow up: 39 weeks; assessed with: SF-12-mental)												
1	randomised trials	very serious [1]	not serious	not serious	very serious [3]	none	14	-	-	SMD **0.26 fewer** (1.03 fewer to 0.51 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Carer mental health (follow up: 39 weeks; assessed with: General Health Questionnaire-30)												
1	randomised trials	very serious [1]	not serious	not serious	very serious [3]	none	14	-	-	SMD **0.11 fewer** (0.88 fewer to 0.66 more)	⨁◯◯◯ VERY LOW	IMPORTANT

1. Risk of selective outcome, performance, and detection bias
2. Confidence intervals cross one minimally important difference. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).
3. Confidence intervals cross two minimally important differences. Sample size less than optimal information size (<400 for continuous outcomes or <300 for dichotomous outcomes).

N.6 Interventions to enhance carer well-being

N.6.1 Interventions informed by cognitive behavioural principles versus control for family carers

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cognitive behavioural intervention	any control	Relative (95% CI)	Absolute (95% CI)		
Carer health and well-being (depression) – post-treatment												
5	randomised trials	serious [1]	not serious	serious [2]	serious [3]	none	251	-	-	SMD **0.35 fewer** (0.54 fewer to 0.15 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (depression) – follow-up (follow up: range 46 to 104 weeks to)												
2	randomised trials	serious [1]	not serious	serious [2]	serious [3]	none	64	-	-	SMD **0.41 fewer** (0.79 fewer to 0.04 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (clinically depressed) – post-treatment												
1	randomised trials	serious [1]	not serious	serious [2]	very serious [4]	none	3/53 (5.7%)	13/58 (22.4%)	RR 0.25 (0.08 to 0.84)	**168 fewer per 1000** (from 36 fewer to 206 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (anxiety, trait) – post-treatment												
2	randomised trials	serious [1]	not serious	serious [2]	serious [3]	none	37	-	-	SMD **0.5 fewer** (1.03 fewer to 0.03 more)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (anxiety, state) – post-treatment												
1	randomised trials	serious [5]	not serious	serious [2]	very serious [4]	none	18	-	-	SMD **0.46 fewer** (1.12 fewer to 0.2 more)	⨁◯◯◯ VERY LOW	CRITICAL

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cognitive behavioural intervention	any control	Relative (95% CI)	Absolute (95% CI)		
Carer health and well-being (mental ill health) – post-treatment												
1	randomised trials	serious [5]	not serious	serious [2]	very serious [4]	none	29	-	-	SMD **2.19 fewer** (2.85 fewer to 1.53 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (quality of life) – post-treatment												
1	randomised trials	serious [5]	not serious	serious [2]	very serious [4]	none	29	-	-	SMD **0.87 more** (0.33 more to 1.41 more)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (stress) – post-treatment												
3	randomised trials	serious [1]	serious [6]	serious [2]	serious [3]	none	225	-	-	SMD **0.45 fewer** (0.78 fewer to 0.12 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (stress) – follow-up (follow up: mean 104 weeks)												
1	randomised trials	serious [5]	not serious	serious [2]	very serious [4]	none	49	-	-	SMD **0.43 fewer** (0.9 fewer to 0.05 more)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (clinically stressed) – post-treatment												
1	randomised trials	serious [5]	not serious	serious [2]	very serious [4]	none	2/53 (3.8%)	17/58 (29.3%)	RR 0.13 (0.03 to 0.53)	**255 fewer per 1000** (from 138 fewer to 284 fewer)	⨁◯◯◯ VERY LOW	CRITICAL

1. Most information is from studies at moderate risk of bias
2. Population not family carers of people with learning disabilities with no mental health problems.
3. Optimal information size not met
4. Optimal information size not met; small, single study

5. Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
6. I2 > 40%. This is the criterion that was used in the challenging behaviour guideline.

N.6.2 Psychosocial support interventions versus control for parents

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Psychosocial support interventions	any control	Relative (95% CI)	Absolute (95% CI)		
Carer health and well-being (stress) – post-treatment												
1	randomised trials	serious [1]	not serious	serious [2]	very serious [3]	none	16	-	-	SMD **1.21 fewer** (2.04 fewer to 0.39 fewer)	⨁◯◯◯ VERY LOW	CRITICAL

1. Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
2. Population not family carers of people with learning disabilities with no mental health problems.
3. Optimal information size not met; small, single study

N.6.3 Psychoeducation versus control for parents

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Psychoeducation	any control	Relative (95% CI)	Absolute (95% CI)		
Carer health and well-being (depression) – follow-up (follow up: mean 4 weeks)												
1	randomised trials	serious [1]	not serious	serious [2]	very serious [3]	none	40	-	-	SMD **0.84 fewer** (1.31 fewer to 0.36 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (burnout) – follow-up (follow up: mean 8 weeks)												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Psychoeducation	any control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious [1]	not serious	serious [2]	very serious [3]	none	45	-	-	SMD **0.35 fewer** (0.77 fewer to 0.06 more)	⨁◯◯◯ VERY LOW	CRITICAL

1. Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
2. Population not family carers of people with learning disabilities with no mental health problems.
3. Optimal information size not met; small, single study

N.6.4 Mindfulness versus control for staff

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Mindfulness interventions	any control	Relative (95% CI)	Absolute (95% CI)		
Carer health and well-being (mental well-being) – post-treatment												
1	randomised trials	serious [1]	not serious	serious [2]	very serious [3]	none	66	-	-	SMD **0.17 more** (0.19 fewer to 0.53 more)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (mental well-being) – follow-up (follow up: mean 6 weeks)												
1	randomised trials	serious [1]	not serious	serious [2]	very serious [3]	none	66	-	-	SMD **0.28 more** (0.08 fewer to 0.64 more)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (mental ill health) – post-treatment												

Quality assessment							Number of patients		Effect		Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Mindfulness interventions	any control	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	serious [4]	serious [5]	serious [2]	serious [3]	none	84	-	-	SMD **0.54 fewer** (1.06 fewer to 0.02 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (mental ill health) – follow-up (follow up: range 6-13 weeks to)												
2	randomised trials	serious [4]	serious [5]	serious [2]	serious [3]	none	84	-	-	SMD **0.24 fewer** (0.72 fewer to 0.24 more)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (stress) – post-treatment												
1	randomised trials	serious [1]	not serious	serious [2]	very serious [3]	none	66	-	-	SMD **0.17 more** (0.19 fewer to 0.53 more)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (stress) – follow-up (follow up: mean 6 weeks)												
1	randomised trials	serious [1]	not serious	serious [2]	very serious [3]	none	66	-	-	SMD **0.05 fewer** (0.41 fewer to 0.31 more)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (burnout) – post-treatment												
1	randomised trials	serious [1]	not serious	serious [2]	very serious [3]	none	18	-	-	SMD **0.18 fewer** (0.86 fewer to 0.49 more)	⨁◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (burnout) – follow-up (follow up: mean 13 weeks)												
1	randomised trials	serious [1]	not serious	serious [2]	very serious [3]	none	18	-	-	SMD **0.08 fewer** (0.76 fewer to 0.59 more)	⨁◯◯◯ VERY LOW	CRITICAL

1. Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
2. Population not family carers of people with learning disabilities with no mental health problems.
3. Optimal information size not met; small, single study
4. Most information is from studies at moderate risk of bias
5. I2 > 40%. This is the criterion that was used in the challenging behaviour guideline.

N.6.5 Mindfulness versus control for parents

Quality assessment							Impact	Quality	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Carer health and well-being (mental well-being) – post-treatment (follow up: 8 weeks; assessed with: CES-D Total depression score)									
1	randomised trials	serious [1]	not serious	serious [2]	serious [3]	none	Parent depression appeared to decrease in the intervention group from baseline (from 17.86 to 11.67) and increase after treatment in the control group from baseline (from 17.53 to 22.0). (no variance reported)	⊕◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (mental ill health) – post-treatment (follow up: 8 weeks; assessed with: PSI Parental Distress Subscale)									
1	randomised trials	serious [1]	not serious	serious [2]	serious [3]	none	Parent distress appeared to decrease in the intervention group from baseline (from 35.17 to 31.72) and also in the control group from baseline (from 38.28 to 37.61). However, the control group appeared to have higher distress at baseline. (no variance reported)	⊕◯◯◯ VERY LOW	CRITICAL
Carer health and well-being (satisfaction with life) – post-treatment (follow up: 8 weeks)									
1	randomised trials	serious [1]	not serious	serious [2]	serious [3]	none	Satisfaction with life appeared to increase in both groups but the increased appeared larger in the intervention group (19.8 to 24.65 in the intervention group versus from 18.41 to 19.42 in the control group). (no variance reported)	⊕◯◯◯ VERY LOW	CRITICAL

1. Risk of selection, selective outcomes bias.
2. Population not family carers of people with learning disabilities with no mental health problems.

4. Optimal information size not met; small, single study

N.6.6 Carer outcomes from parent training for child mental health

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Parent training	Waiting list control	Relative (95% CI)	Absolute (95% CI)		
Mental health after individual training (end of treatment) (follow up: range 10 weeks to 16 weeks; assessed with: Depression Anxiety and Stress Scales (DASS))												
2	randomised trials	very serious [1]	not serious	serious [2]	not serious	none	73	-	-	SMD **0.36 SD lower** (1.27 lower to 0.55 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Carer satisfaction after individual training (end of treatment) (follow up: range 10 weeks to 16 weeks; assessed with: Parenting Sense of Competence Scale (PSOC))												
1	randomised trials	not serious	not serious	serious [2]	serious [3]	none	50	-	-	SMD **0.81 SD higher** (0.3 higher to 1.31 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life after individual training (end of treatment) (follow up: range 10 weeks to 16 weeks; assessed with: Abbreviated Dyadic Adjustment Scale (ADAS))												
1	randomised trials	not serious	not serious	serious [2]	serious [3]	none	50	-	-	SMD **0.29 SD higher** (0.2 lower to 0.78 higher)	⨁⨁◯◯ LOW	CRITICAL
Stress after individual parent training (end of treatment) (follow up: range 10 weeks to 16 weeks; assessed with: Parenting Scale)												
1	randomised trials	not serious	not serious	serious [2]	serious [3]	none	50	-	-	SMD **0.55 SD lower** (1.05 lower to 0.05 lower)	⨁⨁◯◯ LOW	IMPORTANT
Mental health after standard or enhanced individual parent training (follow up: mean 52 weeks; assessed with: Depression Anxiety and Stress Scales (DASS))												

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Parent training	Waiting list control	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	serious [2]	serious [3]	none	23	19	-	MD **5.98 lower** (15.13 lower to 3.17 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life after standard or enhanced individual parent training (follow up: mean 52 weeks; assessed with: Abbreviated Dyadic Adjustment Scale (ADAS))												
1	randomised trials	not serious	not serious	serious [2]	serious [3]	none	19	23	-	MD **0.73 higher** (1.95 lower to 3.41 higher)	⨁⨁◯◯ LOW	CRITICAL
Carer satisfaction after standard or enhanced individual parent training (follow up: mean 52 weeks; assessed with: Parenting Sense of Competence Scale (PSOC))												
1	randomised trials	not serious	not serious	serious [2]	serious [3]	none	19	23	-	MD **0.43 higher** (7.27 lower to 8.13 higher)	⨁⨁◯◯ LOW	CRITICAL
Stress after standard or enhanced individual parent training (follow up: mean 52 weeks; assessed with: Parenting Scale)												
1	randomised trials	not serious	not serious	serious [2]	serious [3]	none	23	19	-	MD **0.15 higher** (0.23 lower to 0.53 higher)	⨁⨁◯◯ LOW	IMPORTANT
Carer satisfaction after group parent training (end of treatment) (follow up: mean 8 weeks; assessed with: Kansas Parental Satisfaction Scale - Short Form (KPS-SF))												
1	randomised trials	serious [4]	not serious	serious [2]	serious [3]	none	16	13	-	MD **3.43 higher** (0.54 higher to 6.32 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Stress after group parent training (follow up: mean 8 weeks; assessed with: Parenting Stress Index (short and long forms))												

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Parent training	Waiting list control	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	very serious [5]	serious [6]	serious [2]	serious [3]	none	30	-	-	SMD **0.08 SD higher** (0.44 lower to 0.61 higher)	⨁◯◯◯ VERY LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference

1. Downgraded as high risk of bias on allocation concealment, missing outcome data and unclear risk of selective reporting
2. Downgraded as patients have learning disabilities but no mental health problem
3. Downgraded as small sample size
4. Downgraded as high risk of performance and detection bias
5. Downgraded for unclear allocation concealment and high risk of performance and detection bias
6. Downgraded as studies show opposing direction of effect