Appendix O: Clinical evidence – GRADE evidence profiles for all studies

Contents_Toc418860862

A.1 Risk markers associated with the development of behaviour that challenges

A.1.1 Auditory impairment

Table O.1: Auditory impairment versus no auditory impairment as a risk factor for challenging behaviour

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With no impairment	With auditory impairment		Risk with no impairment	Risk difference with auditory impairment (95% CI)
All aggression (physical, verbal and destructive) (assessed with: Validated questionnaire)											
1938 (2 studies)	no serious risk of bias	serious[1]	no serious indirectness	no serious imprecision	undetected	⨁⊖⊖⊖ VERY LOW[1] due to inconsistency	380/1628 (23.3%)	35/310 (11.3%)	OR 0.97 (0.42 to 2.23)	233 per 1000	5 fewer per 1000 (from 120 fewer to 171 more)
Self-injury (assessed with: Validated questionnaire)											
2086 (3 studies)	no serious risk of bias	serious[1]	no serious indirectness	no serious imprecision	undetected	⨁⊖⊖⊖ VERY LOW[1] due to inconsistency	419/1770 (23.7%)	37/316 (11.7%)	OR 1.05 (0.49 to 2.29)	237 per 1000	9 more per 1000 (from 105 fewer to 179 more)
Stereotypy (assessed with: Validated questionnaire)											
915 (1 study)	no serious risk	no serious inconsistency	no serious indirectness	serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[2] due to	362/881 (41.1%)	16/34 (47.1%)	OR 1.27 (0.64 to 2.53)	411 per 1000	59 more per 1000 (from 102

Quality assessment							Summary of findings				
	of bias					imprecision					fewer to 227 more)

[1] I^2 > 40%
[2] Optimal information size not met; single study

A.2 Autism diagnosis

Table O.2: Autism diagnosis versus no autism diagnosis as a risk factor for challenging behaviour

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With no autism diagnosis	With autism diagnosis		Risk with no autism diagnosis	Risk difference with autism diagnosis (95% CI)
All aggression (physical, verbal and destructive) (assessed with: Validated questionnaires, interviews and medical records)											
1938 (2 studies)	no serious risk of bias	serious[1]	no serious indirectness	no serious imprecision	undetected	⨁⊖⊖⊖ VERY LOW[1] due to inconsistency	337/1718 (19.6%)	78/220 (35.5%)	OR 1.76 (1.17 to 2.65)	196 per 1000	104 more per 1000 (from 26 more to 197 more)
Destruction of property (assessed with: Questionnaire and interviews with both service user and carer)											
2376 (2 studies)	no serious risk of bias	very serious[2]	no serious indirectness	no serious imprecision	undetected	⨁⊖⊖⊖ VERY LOW[2,3] due to inconsistency, large effect	121/1285 (9.4%)	279/1091 (25.6%)	OR 5.6 (1.39 to 22.56)	94 per 1000	274 more per 1000 (from 32 more to 607 more)

Quality assessment							Summary of findings				
Physical aggression (assessed with: Validated questionnaires, interviews and medical records)											
5637 (4 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	undetected	⨁⨁⨁⊖ MODERATE[3] due to large effect	712/4468 (15.9%)	357/1169 (30.5%)	RR 2.80 (1.98 to 3.98)	159 per 1000	287 more per 1000 (from 156 more to 475 more)
Self-injury (assessed with: Validated questionnaires and interviews with both service user and carer)											
4338 (5 studies)	no serious risk of bias	very serious[2]	no serious indirectness	no serious imprecision	undetected	⨁⊖⊖⊖ VERY LOW[2,3] due to inconsistency, large effect	416/3015 (13.8%)	390/1323 (29.5%)	OR 3.11 (1.81 to 5.35)	138 per 1000	194 more per 1000 (from 87 more to 323 more)

[1] I^2 > 40%
[2] I^2 > 75%
[3] RR >2

A.2.1 Degree of learning disability

Table O.3: Mild/moderate learning disability versus severe/profound learning disability as a risk factor for challenging behaviour

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With Mild/ moderate LD	With Severe/profound LD		Risk with Mild/ moderate LD	Risk difference with Severe/profound LD (95% CI)
All aggression (physical, verbal and destructive) (assessed with: Validated questionnaires)											
1918 (2 studies)	no serio	very serious[1]	no serious indirectnes	no serious	undetected	⨁⊖⊖⊖ VERY LOW[1]	300/1398 (21.5%)	111/520 (21.3%)	OR 1.70	215 per 1000	103 more per 1000

Quality assessment							Summary of findings				
	us risk of bias		s	imprecisio n		due to inconsistenc y			(0.81 to 3.57)		(from 33 fewer to 279 more)
Challenging behaviour (global) (assessed with: Survey)											
822 (1 study)	no serio us risk of bias	no serious inconsistenc y	no serious indirectnes s	serious[2]	undetecte d	⨁⨁⨀⨀ LOW[2,3] due to imprecision, large effect	40/604 (6.6%)	51/218 (23.4%)	OR 4.31 (2.75 to 6.74)	66 per 1000	168 more per 1000 (from 97 more to 257 more)
Destruction of property (assessed with: Validated questionnaire)											
3160 (1 study) 12 months	no serio us risk of bias	no serious inconsistenc y	no serious indirectnes s	serious[2]	undetecte d	⨁⨀⨀⨀ VERY LOW[2] due to imprecision	496/2165 (22.9%)	259/995 (26%)	OR 1.18 (1 to 1.41)	229 per 1000	31 more per 1000 (from 0 more to 66 more)
Inappropriate sexual behaviour (assessed with: Validated questionnaire)											
3160 (1 study) 12 months	no serio us risk of bias	no serious inconsistenc y	no serious indirectnes s	serious[2]	undetecte d	⨁⨀⨀⨀ VERY LOW[2] due to imprecision	211/2165 (9.7%)	99/995 (9.9%)	OR 1.02 (0.8 to 1.32)	97 per 1000	2 more per 1000 (from 18 fewer to 27 more)
Physical aggression – inpatient setting (assessed with: Survey)											
11139 (1 study)	no serio us risk of bias	no serious inconsistenc y	serious[4]	serious[2]	undetecte d	⨁⨀⨀⨀ VERY LOW[2,4] due to indirectness, imprecision	731/2485 (29.4%)	1885/8654 (21.8%)	OR 0.67 (0.6 to 0.74)	294 per 1000	76 fewer per 1000 (from 58 fewer to 94 fewer)
Physical aggression – mixed setting (assessed with: Validated questionnaires, interviews, observations and medical records)											
43864 (6 studies)	no serio us	very serious[1]	no serious indirectnes s	no serious imprecisio	undetecte d	⨁⨀⨀⨀ VERY LOW[1] due to	2831/2079 4 (13.6%)	4189/23070 (18.2%)	OR 1.76 (1.4 to	136 per 1000	81 more per 1000 (from 45 more

Quality assessment							Summary of findings				
	risk of bias			n		inconsistency			2.2)		to 121 more)
Self-injury (assessed with: Validated questionnaires, surveys and medical records)											
85888 (12 studies) 0 to 36 months	no serious risk of bias	very serious[1]	no serious indirectness	no serious imprecision	undetected	⨁⊖⊖⊖ VERY LOW[1,3] due to inconsistency, large effect	2144/40811 (5.3%)	7584/45077 (16.8%)	OR 3.75 (2.62 to 5.38)	53 per 1000	120 more per 1000 (from 74 more to 177 more)
Stereotypy (assessed with: Validated questionnaires and surveys)											
23946 (4 studies)	no serious risk of bias	very serious[1]	no serious indirectness	no serious imprecision	undetected	⨁⊖⊖⊖ VERY LOW[1,3] due to inconsistency, large effect	1153/17847 (6.5%)	2740/6099 (44.9%)	OR 6.38 (1.42 to 28.65)	65 per 1000	241 more per 1000 (from 25 more to 600 more)
Verbal aggression (assessed with: Validated questionnaire)											
3160 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[2] due to imprecision	896/2165 (41.4%)	293/995 (29.4%)	OR 0.59 (0.5 to 0.69)	414 per 1000	120 fewer per 1000 (from 86 fewer to 153 fewer)

[1] I^2 > 75%
[2] Optimal information size not met; single study
[3] RR > 2
[4] Partial applicability to review population – high risk inpatient

A.2.2 Expressive communication difficulties

Table O.4: Expressive communication difficulties versus no expressive communication difficulties as a risk factor for challenging behaviour

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With no expressive communication difficulties	With expressive communication difficulties		Risk with no expressive communication difficulties	Risk difference with expressive communication difficulties (95% CI)
All aggression (physical, verbal and destructive) (assessed with: Validated questionnaire)											
1936 (2 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	undetected	⨁⨁⨀⨀ LOW	300/1310 (22.9%)	115/626 (18.4%)	OR 1.41 (1.08 to 1.86)	229 per 1000	66 more per 1000 (from 14 more to 127 more)
Physical aggression- adult population (assessed with: Questionnaire)											
3662 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	undetected	⨁⨀⨀⨀ VERY LOW[1,2] due to risk of bias, imprecision	784/2994 (26.2%)	250/668 (37.4%)	OR 1.69 (1.41 to 2.01)	262 per 1000	113 more per 1000 (from 72 more to 154 more)
Physical aggression- mixed population (assessed with: Non-validated questionnaire)											
211 (1 study)	serious[3]	no serious inconsistency	no serious indirectness	serious[2]	undetected	⨁⨁⨀⨀ LOW[2,3,4] due to risk of bias, imprecision,	52/166 (31.3%)	2/45 (4.4%)	OR 0.10 (0.02 to 0.44)	313 per 1000	270 fewer per 1000 (from 146 fewer to 304 fewer)

Quality assessment							Summary of findings				
						large effect					
Self-injury (assessed with: Questionnaires, interviews and formal assessments)											
7502 (9 studies) 0 to 3 years	no serious risk of bias	very serious[5]	no serious indirectness	no serious imprecision	undetected	⨁⊖⊖⊖ VERY LOW[5,6] due to inconsistency, large effect	821/5630 (14.6%)	566/1872 (30.2%)	OR 2.93 (1.8 to 4.78)	146 per 1000	188 more per 1000 (from 89 more to 304 more)
Stereotypy (assessed with: Validated questionnaire)											
915 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[2] due to imprecision	290/769 (37.7%)	88/146 (60.3%)	OR 2.51 (1.74 to 3.6)	377 per 1000	226 more per 1000 (from 136 more to 308 more)

1 Non validated checklist for risk and outcome assessment
[2] Optimal information size not met; single study
3 Questionnaire for risk and outcome assessment was not validated
[4] RR < 0.2
[5] I^2 > 75%
[6] RR > 2

A.2.4 Receptive communication difficulties

Table O.5: Receptive communication difficulties versus no receptive communication difficulties as a risk factor for challenging behaviour

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With no receptive communication difficulties	With receptive communication difficulties		Risk with no receptive communication difficulties	Risk difference with receptive communication difficulties (95% CI)
Self-injury (assessed with: Questionnaire and interview)											
1321 (3 studies) 0 to 3 years	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	undetected	⨁⨁⨁⊖ MODERATE[1] due to large effect	148/1098 (13.5%)	82/223 (36.8%)	OR 3.46 (2.5 to 4.79)	135 per 1000	215 more per 1000 (from 146 more to 293 more)

[1] RR > 2

A.2.5 Gender

Table O.6: Male gender versus female gender as a risk factor for challenging behaviour

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With female gender	With male gender		Risk with female gender	Risk difference with male gender (95% CI)
All aggression (physical, verbal and destructive) (assessed with: Validated questionnaire and observation)											
2046 (3 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	undetected	⨁⨁⊖⊖ LOW	237/898 (26.4%)	238/1148 (20.7%)	OR 0.63 (0.51 to 0.79)	264 per 1000	80 fewer per 1000 (from 43 fewer to 109 fewer)
Challenging behaviour (global) (assessed with: Validated survey)											
816 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	undetected	⨁⊖⊖⊖ VERY LOW[1] due to imprecision	34/370 (9.2%)	56/446 (12.6%)	OR 1.42 (0.9 to 2.23)	92 per 1000	34 more per 1000 (from 8 fewer to 92 more)
Destruction of property (assessed with: Validated questionnaire)											
3461 (2 studies) 0 to 12 months	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	undetected	⨁⨁⊖⊖ LOW	0/3461 (0%)2	-2	Not estimable	See comment[2]	-
Inappropriate sexual behaviour (assessed with: Questionnaire)											
3160	no	no serious	no serious	serious[1]	undetecte	⨁⊖⊖⊖	116/1527	194/1633	OR 1.64	76 per	43 more

Quality assessment							Summary of findings				
(1 study) 12 months	serious risk of bias	inconsistency	indirectness		d	VERY LOW[1] due to imprecision	(7.6%)	(11.9%)	(1.29 to 2.09)	1000	per 1000 (from 20 more to 71 more)
Physical aggression (assessed with: Validated questionnaires, interviews, observations and medical records)											
6925 (5 studies) 0 to 12 months	no serious risk of bias	serious[3]	no serious indirectness	no serious imprecision	undetected	⊕⊖⊖⊖ VERY LOW[3] due to inconsistency	0/6925 (0%)2	-2	Not estimable	See comment[2]	-
Self-injury – mixed settings (assessed with: Questionnaire and survey)											
6174 (6 studies) 0 to 12 months	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	undetected	⊕⊕⊖⊖ LOW	827/2820 (29.3%)	879/3354 (26.2%)	OR 0.81 (0.69 to 0.96)	293 per 1000	42 fewer per 1000 (from 8 fewer to 71 fewer)
Self-injury- inpatient setting (assessed with: Non-validated questionnaire, survey and interview)											
18227 (5 studies) 0 to 3 years	no serious risk of bias	very serious[4]	no serious indirectness	no serious imprecision	undetected	⊕⊖⊖⊖ VERY LOW[4] due to inconsistency	1008/8246 (12.2%)	1220/9981 (12.2%)	OR 0.97 (0.76 to 1.23)	122 per 1000	3 fewer per 1000 (from 27 fewer to 24 more)
Stereotypy (assessed with: Validated questionnaire)											
915 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	undetected	⊕⊖⊖⊖ VERY LOW[1] due to imprecision	169/411 (41.1%)	209/504 (41.5%)	RR 1.01 (0.86 to 1.18)	411 per 1000	4 more per 1000 (from 58 fewer to 74 more)
Verbal aggression (assessed with: Validated questionnaire)											
3461 (2 studies) 0 to 12 months	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	undetected	See comment	0/3461 (0%)2	-2	Not estimable	See comment[2]	-

Quality assessment	Summary of findings
[1] Optimal information size not met; single study [2] N/A; Generic inverse variance [3] I^2 > 40% [4] I^2 > 75%	

A.2.6 Mental health needs

Table O.7: Mental health needs versus no mental health needs as a risk factor for challenging behaviour

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With no mental health needs	With mental health needs		Risk with no mental health needs	Risk difference with mental health needs (95% CI)
All aggression (physical, verbal and destructive) (assessed with: Validated questionnaire)											
1938 (2 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	undetected	⨁⨁⊖⊖ LOW	377/1837 (20.5%)	38/101 (37.6%)	OR 2.03 (1.3 to 3.15)	205 per 1000	139 more per 1000 (from 46 more to 243 more)
Destruction of property (assessed with: Validated questionnaire and survey)											
30874 (2 studies)	no serious risk of bias	very serious[1]	no serious indirectness	no serious imprecision	undetected	⨁⊖⊖⊖ VERY LOW[1] due to inconsistency	0/30874 (0%)2	-2	Not estimable	See comment[2]	-

Quality assessment							Summary of findings				
Physical aggression (assessed with: Validated questionnaire and survey)											
30874 (2 studies)	no serious risk of bias	very serious[1]	no serious indirectness	no serious imprecision	undetected	⨁⊖⊖⊖ VERY LOW[1] due to inconsistency	0/30874 (0%)2	-2	Not estimable	See comment[2]	-
Self-injury (assessed with: Validated questionnaires and survey)											
32516 (3 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	undetected	⨁⨁⊖⊖ LOW	2690/28860 (9.3%)	450/3656 (12.3%)	OR 1.4 (1.26 to 1.56)	93 per 1000	33 more per 1000 (from 21 more to 45 more)
Stereotypy (assessed with: Validated questionnaire and survey)											
31493 (2 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	undetected	⨁⨁⊖⊖ LOW	1970/27876 (7.1%)	293/3617 (8.1%)	OR 1.26 (1.1 to 1.43)	71 per 1000	17 more per 1000 (from 7 more to 27 more)
Verbal aggression (assessed with: Validated questionnaire and survey)											
30874 (2 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	undetected	⨁⨁⨁⊖ MODERATE[3] due to large effect	0/30874 (0%)2	-2	Not estimable	See comment[2]	-

[1] I^2 > 75%
[2] N/A; Generic inverse variance
[3] RR > 2

A.2.7 Mobility impairment

Table O.8: Mobility impairment versus no mobility impairment as a risk factor for challenging behaviour

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With no impairment	With mobility impairment		Risk with no impairment	Risk difference with mobility impairment (95% CI)
All aggression (physical, verbal and destructive) (assessed with: Validated questionnaire)											
1023 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	undetected	⨁⊖⊖⊖ VERY LOW[1] due to imprecision	78/775 (10.1%)	22/248 (8.9%)	OR 0.87 (0.53 to 1.43)	101 per 1000	12 fewer per 1000 (from 45 fewer to 37 more)
Self-injury- adult population (assessed with: Validated questionnaire)											
1023 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	undetected	⨁⊖⊖⊖ VERY LOW[1] due to imprecision	78/775 (10.1%)	22/248 (8.9%)	OR 0.87 (0.53 to 1.43)	101 per 1000	12 fewer per 1000 (from 45 fewer to 37 more)
Self-injury- children and young people population (assessed with: Validated questionnaire)											
147 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	undetected	⨁⊖⊖⊖ VERY LOW[1] due to imprecision	64/134 (47.8%)	9/13 (69.2%)	OR 2.46 (0.72 to 8.38)	478 per 1000	215 more per 1000 (from 81 fewer to 407 more)

[1] Optimal information size not met; single study

A.2.8 Visual impairment

Table O.9: Visual impairment versus no visual impairment as a risk factor for challenging behaviour

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With No impairment	With Visual impairment		Risk with No impairment	Risk difference with Visual impairment (95% CI)
All aggression (physical, verbal and destructive) (assessed with: Validated questionnaire)											
1938 (2 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	undetected	⨁⨁⊖⊖ LOW	349/1422 (24.5%)	66/516 (12.8%)	OR 1.22 (0.78 to 1.92)	245 per 1000	39 more per 1000 (from 43 fewer to 139 more)
Self-injury (assessed with: Validated questionnaire)											
2086 (3 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	undetected	⨁⨁⊖⊖ LOW	384/1564 (24.6%)	73/522 (14%)	OR 1.45 (1.02 to 2.06)	246 per 1000	75 more per 1000 (from 4 more to 156 more)
Stereotypy (assessed with: Validated questionnaire)											
915 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	undetected	⨁⊖⊖⊖ VERY LOW[1] due to imprecision	356/880 (40.5%)	22/35 (62.9%)	OR 2.49 (1.24 to 5.01)	405 per 1000	224 more per 1000 (from 53 more to 368 more)

[1] Optimal information size; single study

A.3 Interventions aimed at the prevention of behaviour that challenges

A.3.1 Educational intervention versus attention control

Table O.10: Learning Experiences and Alternative Program for Pre-schoolers and Their Parents (LEAP) – full replication condition versus manual-only attention control

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With attention control	With educational intervention		Risk with attention control	Risk difference with educational intervention (95% CI)
Behaviour that challenges (severity) – post-treatment (measured with: Change score[1]; Better indicated by lower values)											
294 (1 study)	serious[2]	no serious inconsistency	serious[3]	serious[4]	undetected	⨁⊖⊖⊖ VERY LOW[2,3,4] due to risk of bias, indirectness, imprecision	117	177	-		The mean behaviour that challenges (severity) – post-treatment in the intervention groups was 0.19 standard deviations lower (0.42 lower to 0.04 higher)
Adaptive functioning (social) – post-treatment (Better indicated by lower values)											
294 (1 study)	serious[2]	no serious inconsistency	serious[3]	serious[4]	undetected	⨁⊖⊖⊖ VERY LOW[2,3,4] due to risk of bias,	117	177	-		The mean adaptive functioning (social) – post-treatment in the

Quality assessment							Summary of findings				
						indirectness, imprecision					intervention groups was 0.76 standard deviations higher (0.52 to 1 higher)
Adaptive functioning (communication) – post-treatment (Better indicated by lower values)											
294 (1 study)	serious[2]	no serious inconsistency	serious[3]	serious[4]	undetected	⨁⊖⊖⊖ VERY LOW[2,3,4] due to risk of bias, indirectness, imprecision	117	177	-		The mean adaptive functioning (communication) – post-treatment in the intervention groups was 0.94 standard deviations higher (0.7 to 1.19 higher)

[1] Due to significant baseline differences, standard deviation of change and estimates of mean change were derived using initial and final mean values and utilising r = 0.5. Sensitivity analyses were used to explore the impact of altering assumptions about the calculation of the effect size, but this resulted in no change to conclusions.
[2] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect

A.3.2 Home-based EBI versus centre-based EBI

Table O.11: Home-based Building Blocks programme versus centre-based Building Blocks programme

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With centre-based early behavioural intervention	With home-based early behavioural intervention		Risk with centre-based early behavioural intervention	Risk difference with home-based early behavioural intervention (95% CI)
Behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
44 (1 study)	serious [1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	22	22	-		The mean behaviour that challenges (severity) – post-treatment in the intervention groups was 0.11 standard deviations lower (0.7 lower to 0.48 higher)
Adaptive functioning (social) – post-treatment (Better indicated by lower values)											
56 (1 study)	serious [1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecisio	29	27	-		The mean adaptive functioning (social) – post-treatment in the intervention

Quality assessment							Summary of findings				
						n					groups was 0.63 standard deviations lower (1.17 to 0.09 lower)
Adaptive functioning (communication) – post-treatment (Better indicated by lower values)											
55 (1 study)	serious [1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	29	26	-		The mean adaptive functioning (communication) – post-treatment in the intervention groups was 0.46 standard deviations lower (1 lower to 0.07 higher)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Optimal information size not met; small, single study

A.3.3 EIBI versus parent training

Table O.12: EIBI (UCLA model) versus parent training

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With parent training	With EIBI		Risk with parent training	Risk difference with EIBI (95% CI)

Quality assessment							Summary of findings				
Behaviour that challenges (severity) – post-treatment (measured with: Parent-rated; Better indicated by lower values)											
28 (1 study)	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to indirectness, imprecision	13	15	-		The mean behaviour that challenges (severity) – post-treatment in the intervention groups was 0.36 standard deviations lower (1.1 lower to 0.39 higher)
Behaviour that challenges (severity) – post-treatment (measured with: Teacher-report; Better indicated by lower values)											
28 (1 study)	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to indirectness, imprecision	13	15	-		The mean behaviour that challenges (severity) – post-treatment in the intervention groups was 0.47 standard deviations higher (0.28 lower to 1.23 higher)
Adaptive functioning (communication) – post-treatment (Better indicated by lower values)											
28 (1 study)	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to indirectness, imprecision	13	15	-		The mean adaptive functioning (communication) – post-treatment in the intervention groups was 0.63 standard deviations higher (0.13 lower to 1.39 higher)

Quality assessment							Summary of findings				
Adaptive functioning (global) – post-treatment (Better indicated by lower values)											
28 (1 study)	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to indirectness, imprecision	13	15	-		The mean adaptive functioning (global) – post-treatment in the intervention groups was 0.11 standard deviations higher (0.64 lower to 0.85 higher)

[1] Applicability concerns: autism population; no information reported concerning learning disability
[2] Optimal information size not met; small, single study

A.3.4 High supervision EIBI versus low supervision EIBI

Table O.13: High supervision EIBI (clinic-directed UCLA model) versus low supervision EIBI (parent-directed UCLA model)

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With low supervision EIBI (parent-directed)	With high supervision EIBI (clinic-directed)		Risk with low supervision EIBI (parent-directed)	Risk difference with high supervision EIBI (clinic-directed) (95% CI)
Adaptive functioning (communication) -post-treatment (Better indicated by lower values)											
23 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk	10	13	-		The mean adaptive functioning (communication

Quality assessment							Summary of findings				
						of bias, indirectness, imprecision					) -post-treatment in the intervention groups was 0.25 standard deviations lower (1.08 lower to 0.57 higher)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Applicability concerns: autism population; no information reported concerning learning disability
[3] Optimal information size not met; small, single study

A.3.5 Parent training versus any control

Table O.14: Parent training (plus centre based EBI) versus treatment as usual (centre-based EBI)

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With control	With parent training		Risk with control	Risk difference with parent training (95% CI)
Behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
57 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⊖⊖ LOW[1] due to imprecision	28	29	-		The mean behaviour that challenges (severity) – post-treatment in the intervention groups

Quality assessment							Summary of findings				
											was 0.4 standard deviations lower (0.93 lower to 0.12 higher)
Behaviour that challenges (severity) – follow up (Better indicated by lower values)											
117 (2 studies) 26 to 52 weeks	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	undetected	⨁⨁⨀⨀ LOW[2,3] due to risk of bias, imprecision	58	59	-		The mean behaviour that challenges (severity) – follow up in the intervention groups was 0.37 standard deviations lower (0.79 lower to 0.05 higher)
Adaptive functioning (global) – post-treatment (Better indicated by lower values)											
58 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecision	28	30	-		The mean adaptive functioning (global) – post-treatment in the intervention groups was 0.25 standard deviations higher (0.27 lower to 0.77 higher)
Adaptive functioning (global) – follow-up (Better indicated by lower values)											
119 (2 studies) 26 to 52 weeks	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	undetected	⨁⨁⨀⨀ LOW[2,3] due to risk of bias, imprecision	56	63	-		The mean adaptive functioning (global) – follow-up in the intervention groups was 0.52 standard deviations higher

Quality assessment							Summary of findings				
											(0.15 to 0.88 higher)
Adaptive functioning (communication) – follow-up (Better indicated by lower values)											
68 (1 study) 26 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecision	33	35	-		The mean adaptive functioning (communication) – follow-up in the intervention groups was 0.75 standard deviations higher (0.26 to 1.25 higher)

[1] Optimal information size not met; small, single study
2 Most information is from studies at moderate risk of bias
[3] Optimal information size not met

A.4 Interventions aimed at reducing health risks and increasing understanding of physical illness in relation to the prevention or management of behaviour that challenges

A.4.1 Hand-held health record versus treatment as usual

Table O.15: Advocacy Skills Kit Diary or Personal Health Profile versus treatment as usual

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With treatment as usual	With hand-held health record		Risk with treatment as usual	Risk difference with hand-held health record (95% CI)

Quality assessment							Summary of findings				
Health promotion (blood pressure checked)											
119 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⊖⊖ LOW[1] due to imprecision	32/68 (47.1%)	28/51 (54.9%)	RR 1.17 (0.82 to 1.66)	471 per 1000	80 more per 1000 (from 85 fewer to 311 more)
Health promotion (constipation investigation)											
119 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⊖⊖ LOW[1] due to imprecision	1/68 (1.5%)	5/51 (9.8%)	RR 6.67 (0.8 to 55.33)	15 per 1000	83 more per 1000 (from 3 fewer to 799 more)
Health promotion (hearing test)											
119 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⊖⊖ LOW[1] due to imprecision	2/68 (2.9%)	3/51 (5.9%)	RR 2 (0.35 to 11.53)	29 per 1000	29 more per 1000 (from 19 fewer to 310 more)
Health promotion (vision test)											
119 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⊖⊖ LOW[1] due to imprecision	4/68 (5.9%)	7/51 (13.7%)	RR 2.33 (0.72 to 7.55)	59 per 1000	78 more per 1000 (from 16 fewer to 385 more)
Health promotion (weight measured)											
119 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⊖⊖ LOW[1] due to imprecision	17/68 (25%)	18/51 (35.3%)	RR 1.41 (0.81 to 2.46)	250 per 1000	102 more per 1000 (from 47 fewer to 365 more)
Health promotion (weight management plan)											
119	no	no serious	no serious	very	undetecte	⨁⨁⊖⊖	12/68	5/51	RR 0.56	176 per	78 fewer per

Quality assessment							Summary of findings				
(1 study) 52 weeks	serious risk of bias	inconsistency	indirectness	serious[1]	d	LOW[1] due to imprecision	(17.6%)	(9.8%)	(0.21 to 1.48)	1000	1000 (from 139 fewer to 85 more)
Health promotion (epilepsy review)											
119 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⊖⊖ LOW[1] due to imprecision	8/68 (11.8%)	11/51 (21.6%)	RR 1.83 (0.8 to 4.23)	118 per 1000	98 more per 1000 (from 24 fewer to 380 more)
Service user knowledge of health problems (measured with: Knowledge of Health Problems and Terminology Checklist (unvalidated measure); Better indicated by higher values)											
66 (1 study)	serious[2]	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	32	34	-		The mean service user knowledge of health problems in the intervention groups was 0.32 standard deviations lower (0.81 lower to 0.16 higher)
Carer knowledge of health problems (measured with: Knowledge of Health Problems and Terminology Checklist (unvalidated measure); Better indicated by higher values)											
144 (1 study)	serious[2]	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	74	70	-		The mean carer knowledge of health problems in the

Quality assessment							Summary of findings				
											intervention groups was 0 standard deviations higher (0.33 lower to 0.33 higher)
Carer satisfaction (Better indicated by lower values)											
101 (1 study)	serious[2]	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	52	49	-		The mean carer satisfaction in the intervention groups was 0 standard deviations higher (0.39 lower to 0.39 higher)
Service user satisfaction (Better indicated by lower values)											
36 (1 study)	serious[2]	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	20	16	-		The mean service user satisfaction in the intervention groups was 0.6 standard deviations higher (0.08 lower to 1.27 higher)
Premature death											
169	serious[2]	no serious	no serious	very	undetecte	⨁⊖⊖⊖	2/88	5/81	RR 2.72	23 per	39 more per

Quality assessment							Summary of findings				
(1 study)		inconsistency	indirectness	serious[1]	d	VERY LOW[1,2] due to risk of bias, imprecision	(2.3%)	(6.2%)	(0.54 to 13.61)	1000	1000 (from 10 fewer to 287 more)

[1] Optimal information size not met; small, single study
[2] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect

A.4.2 Annual health check versus treatment as usual

Table O.16: Comprehensive Health Assessment Program versus treatment as usual

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With treatment as usual	With annual health check		Risk with treatment as usual	Risk difference with annual health check (95% CI)
Health promotion (blood pressure checked)											
574 (2 studies) 52 weeks	no serious risk of bias	very serious[1]	no serious indirectness	serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to inconsistency, imprecision	131/287 (45.6%)	143/287 (49.8%)	RR 1.09 (0.92 to 1.30)	456 per 1000	41 more per 1000 (from 37 fewer to 137 more)
Health promotion (constipation investigation)											
121 (1 study) 52 weeks	no serious risk of	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⨁⊖⊖ LOW[3] due to	1/68 (1.5%)	4/53 (7.5%)	RR 5.13 (0.59 to 44.58)	15 per 1000	61 more per 1000 (from 6

Quality assessment							Summary of findings				
	bias					imprecision					fewer to 641 more)
Health promotion (hearing test)											
574 (2 studies) 52 weeks	no serious risk of bias	serious[4]	no serious indirectness	serious[2]	undetected	⨁⨁⨀⨀ LOW[2,4] due to inconsistency, imprecision	3/287 (1%)	42/287 (14.6%)	RR 12.22 (2.43 to 61.49)	10 per 1000	117 more per 1000 (from 15 more to 632 more)
Health promotion (vision test)											
574 (2 studies) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	undetected	⨁⨁⨁⨀ MODERATE[2] due to imprecision	16/287 (5.6%)	60/287 (20.9%)	RR 3.75 (2.21 to 6.36)	56 per 1000	153 more per 1000 (from 67 more to 299 more)
Health promotion (acuity corrected by glasses)											
453 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⨁⨀⨀ LOW[3] due to imprecision	0/219 (0%)	3/234 (1.3%)	RR 6.55 (0.34 to 126.14)	0 per 1000	-
Health promotion (otoscopic examination)											
453 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⨁⨀⨀ LOW[3] due to imprecision	50/219 (22.8%)	92/234 (39.3%)	RR 1.72 (1.29 to 2.3)	228 per 1000	164 more per 1000 (from 66 more to 297 more)
Health promotion (weight measurement)											
574 (2 studies) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	undetected	⨁⨁⨁⨀ MODERATE[2] due to imprecision	53/287 (18.5%)	129/287 (44.9%)	RR 2.46 (1.87 to 3.23)	185 per 1000	270 more per 1000 (from 161 more to 412 more)

Quality assessment							Summary of findings				
Health promotion (weight management plan)											
574 (2 studies) 52 weeks	no serious risk of bias	serious[4]	no serious indirectness	serious[2]	undetected	⨁⨁⨀⨀ LOW[2,4] due to inconsistency, imprecision	13/287 (4.5%)	22/287 (7.7%)	RR 2.32 (0.66 to 8.14)	45 per 1000	60 more per 1000 (from 15 fewer to 323 more)
Health promotion (epilepsy review)											
121 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⨁⨀⨀ LOW[3] due to imprecision	8/68 (11.8%)	9/53 (17%)	RR 1.44 (0.6 to 3.49)	118 per 1000	52 more per 1000 (from 47 fewer to 293 more)
Identification of physical health problem (hearing loss)											
453 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⨁⨀⨀ LOW[3] due to imprecision	0/219 (0%)	15/234 (6.4%)	RR 29.02 (1.75 to 482.11)	0 per 1000	-
Identification of physical health problem (visual impairment)											
453 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⨁⨀⨀ LOW[3] due to imprecision	1/219 (0.46%)	7/234 (3%)	RR 6.55 (0.81 to 52.82)	5 per 1000	25 more per 1000 (from 1 fewer to 237 more)
Identification of physical health problem (obesity)											
453 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⨁⨀⨀ LOW[3] due to imprecision	4/219 (1.8%)	17/234 (7.3%)	RR 3.98 (1.36 to 11.64)	18 per 1000	54 more per 1000 (from 7 more to 194 more)
Premature death											
453	no	no serious	no serious	very	undetecte	⨁⨁⨀⨀	1/219	1/234	RR 0.94	5 per	0 fewer per

Quality assessment							Summary of findings				
(1 study) 52 weeks	serious risk of bias	inconsistency	indirectness	serious[3]	d	LOW[3] due to imprecision	(0.46%)	(0.43%)	(0.06 to 14.87)	1000	1000 (from 4 fewer to 63 more)

[1] I^2 > 75%
[2] Optimal information size not met
[3] Optimal information size not met; small, single study
[4] I^2 > 40%

A.4.3 Annual health check versus hand-held health record

Table O.17: Comprehensive Health Assessment Program versus Advocacy Skills Kit Diary

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With hand-held health record	With annual health check		Risk with hand-held health record	Risk difference with annual health check (95% CI)
Health promotion (blood pressure checked)											
104 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⊖⊖ LOW[1] due to imprecision	28/51 (54.9%)	26/53 (49.1%)	RR 0.89 (0.62 to 1.29)	549 per 1000	60 fewer per 1000 (from 209 fewer to 159 more)
Health promotion (constipation investigation)											
104 (1 study)	no serious	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⊖⊖ LOW[1]	5/51 (9.8%)	4/53 (7.5%)	RR 0.77 (0.22 to	98 per 1000	23 fewer per 1000

Quality assessment							Summary of findings				
52 weeks	risk of bias					due to imprecision			2.71)		(from 76 fewer to 168 more)
Health promotion (hearing test)											
104 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecision	3/51 (5.9%)	10/53 (18.9%)	RR 3.21 (0.94 to 10.99)	59 per 1000	130 more per 1000 (from 4 fewer to 588 more)
Health promotion (vision test)											
104 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecision	7/51 (13.7%)	11/53 (20.8%)	RR 1.51 (0.64 to 3.60)	137 per 1000	70 more per 1000 (from 49 fewer to 357 more)
Health promotion (weight measured)											
104 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecision	18/51 (35.3%)	29/53 (54.7%)	RR 1.55 (0.99 to 2.42)	353 per 1000	194 more per 1000 (from 4 fewer to 501 more)
Health promotion (weight management plan)											
104 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecision	5/51 (9.8%)	15/53 (28.3%)	RR 2.89 (1.13 to 7.36)	98 per 1000	185 more per 1000 (from 13 more to 624 more)
Health promotion (epilepsy review)											
104 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecision	11/51 (21.6%)	9/53 (17%)	RR 0.79 (0.36 to 1.74)	216 per 1000	45 fewer per 1000 (from 138 fewer to 160 more)

Quality assessment	Summary of findings
[1] Optimal information size not met; small, single study	

A.4.4 Annual health check and hand-held health record versus treatment as usual

Table O.18: Comprehensive Health Assessment Program and Advocacy Skills Kit Diary versus treatment as usual

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With treatment as usual	With annual health check + hand-held health record		Risk with treatment as usual	Risk difference with annual health check + hand-held health record (95% CI)
Health promotion (blood pressure checked)											
138 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⊖⊖ LOW[1] due to imprecision	32/68 (47.1%)	46/70 (65.7%)	RR 1.4 (1.03 to 1.89)	471 per 1000	188 more per 1000 (from 14 more to 419 more)
Health promotion (constipation investigation)											
138 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⊖⊖ LOW[1] due to imprecision	1/68 (1.5%)	4/70 (5.7%)	RR 3.89 (0.45 to 33.89)	15 per 1000	42 more per 1000 (from 8 fewer to 484 more)
Health promotion (hearing test)											
138	no	no serious	no serious	very	undetecte	⨁⨁⊖⊖	2/68	10/70	RR 4.86	29 per	114 more

Quality assessment							Summary of findings				
(1 study) 52 weeks	serious risk of bias	inconsistency	indirectness	serious[1]	d	LOW[1] due to imprecision	(2.9%)	(14.3%)	(1.1 to 21.36)	1000	per 1000 (from 3 more to 599 more)
Health promotion (vision test)											
138 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⊕⊕⊖⊖ LOW[1] due to imprecision	4/68 (5.9%)	20/70 (28.6%)	RR 4.86 (1.75 to 13.47)	59 per 1000	227 more per 1000 (from 44 more to 734 more)
Health promotion (weight measured)											
138 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⊕⊕⊖⊖ LOW[1] due to imprecision	17/68 (25%)	41/70 (58.6%)	RR 2.34 (1.48 to 3.7)	250 per 1000	335 more per 1000 (from 120 more to 675 more)
Health promotion (weight management plan)											
138 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⊕⊕⊖⊖ LOW[1] due to imprecision	12/68 (17.6%)	7/70 (10%)	RR 0.57 (0.24 to 1.35)	176 per 1000	76 fewer per 1000 (from 134 fewer to 62 more)
Health promotion (epilepsy review)											
138 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⊕⊕⊖⊖ LOW[1] due to imprecision	8/68 (11.8%)	7/70 (10%)	RR 0.85 (0.33 to 2.22)	118 per 1000	18 fewer per 1000 (from 79 fewer to 144 more)
[1] Optimal information size not met; small, single study											

A.5 Environmental change interventions aimed at reducing and managing behaviour that challenges

A.5.1 Sensory intervention versus any control

Table O.19: Multisensory room or vibroacoustic chair versus any control

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With any control	With sensory intervention		Risk with any control	Risk difference with sensory intervention (95% CI)
Targeted behaviour that challenges (global) – post-treatment (measured with: Change score[1]; Better indicated by lower values)											
89 (1 study)	serious[2]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[2,3] due to risk of bias, imprecision	41	48	-		The mean targeted behaviour that challenges (global) – post-treatment in the intervention groups was 1.69 standard deviations higher (1.2 to 2.18 higher)
Targeted behaviour that challenges (global) – follow-up (measured with: Change score[1]; Better indicated by lower values)											
89 (1 study) 12 weeks	serious[2]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[2,3]	41	48	-		The mean targeted behaviour that

Quality assessment							Summary of findings				
						due to risk of bias, imprecision					challenges (global) – follow-up in the intervention groups was 0.00 standard deviations higher (0.42 lower to 0.42 higher)
Targeted behaviour that challenges (self-injurious behaviour, severity) – post-treatment (Better indicated by lower values)											
20 (1 study)	serious[2]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[2,3] due to risk of bias, imprecision	10	10	-		The mean targeted behaviour that challenges (self-injurious behaviour, severity) – post-treatment in the intervention groups was 0.2 standard deviations lower (1.08 lower to 0.68 higher)
Targeted behaviour that challenges (self-injurious behaviour, frequency) – post-treatment (Better indicated by lower values)											
20 (1 study)	serious[2]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[2,3] due to risk of bias, imprecision	10	10	-		The mean targeted behaviour that challenges (self-injurious behaviour, frequency) – post-treatment in the intervention

Quality assessment							Summary of findings				
											groups was 0.25 standard deviations lower (1.14 lower to 0.63 higher)
Targeted behaviour that challenges (stereotypical behaviour, severity) – post-treatment (Better indicated by lower values)											
20 (1 study)	serious[2]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[2,3] due to risk of bias, imprecision	10	10	-		The mean targeted behaviour that challenges (stereotypical behaviour, severity) – post-treatment in the intervention groups was 0.33 standard deviations higher (0.55 lower to 1.21 higher)
Targeted behaviour that challenges (stereotypical behaviour, frequency) – post-treatment (Better indicated by lower values)											
20 (1 study)	serious[2]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[2,3] due to risk of bias, imprecision	10	10	-		The mean targeted behaviour that challenges (stereotypical behaviour, frequency) – post-treatment in the intervention groups was 0.22 standard deviations lower (1.1 lower to 0.66

Quality assessment							Summary of findings				
											higher)
Targeted behaviour that challenges (aggressive/ destructive behaviour, severity) – post-treatment (Better indicated by lower values)											
20 (1 study)	serious[2]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[2,3] due to risk of bias, imprecision	10	10	-		The mean targeted behaviour that challenges (aggressive/ destructive behaviour, severity) – post-treatment in the intervention groups was 0.15 standard deviations lower (1.03 lower to 0.72 higher)
Targeted behaviour that challenges (aggressive/ destructive behaviour, frequency) – post-treatment (Better indicated by lower values)											
20 (1 study)	serious[2]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[2,3] due to risk of bias, imprecision	10	10	-		The mean targeted behaviour that challenges (aggressive/ destructive behaviour, frequency) – post-treatment in the intervention groups was 0.22 standard deviations lower (1.1 lower to 0.66 higher)

Quality assessment							Summary of findings				
Adaptive functioning – post-treatment (measured with: Change score[1]; Better indicated by higher values)											
89 (1 study)	serious[2]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[2,3] due to risk of bias, imprecision	41	48	-		The mean adaptive functioning – post-treatment in the intervention groups was 1.12 standard deviations lower (1.57 to 0.67 lower)
Adaptive functioning – follow-up (measured with: Change score[1]; Better indicated by higher values)											
89 (1 study) 12 weeks	serious[2]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[2,3] due to risk of bias, imprecision	41	48	-		The mean adaptive functioning – follow-up in the intervention groups was 0.48 standard deviations lower (0.9 to 0.05 lower)

[1] Due to significant baseline differences, standard deviation of change and estimates of mean change were derived using initial and final mean values and utilising r = 0.5. Sensitivity analyses were used to explore the impact of altering assumptions about the calculation of the effect size, but this resulted in no change to conclusions.
[2] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[3] Optimal information size not met; small, single study

A.5.2 Structured activity versus unstructured activity

Table O.20: Special Olympics Sports Skill Instructional Program versus free play

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With unstructured activity	With structured activity		Risk with unstructured activity	Risk difference with structured activity (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (measured with: Change score[1]; Better indicated by lower values)											
26 (1 study)	serious [2]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[2,3] due to risk of bias, imprecision	13	13	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.87 standard deviations lower (1.68 to 0.06 lower)
Targeted behaviour that challenges (severity) – follow-up (measured with: Change score[1]; Better indicated by lower values)											
26	serious	no serious	no serious	very	undetecte	⨁⊖⊖⊖	13	13	-		The mean

Quality assessment							Summary of findings				
(1 study) 6 weeks	[2]	inconsistency	indirectness	serious[3]	d	VERY LOW[2,3] due to risk of bias, imprecision					targeted behaviour that challenges (severity) – follow-up in the intervention groups was 0.95 standard deviations lower (1.77 to 0.13 lower)

[1] Due to significant baseline differences, standard deviation of change and estimates of mean change were derived using initial and final mean values and utilising r = 0.5. Sensitivity analyses were used to explore the impact of altering assumptions about the calculation of the effect size, but this resulted in no change to conclusions.
[2] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[3] Optimal information size not met; small, single study

A.6 Parent training interventions aimed at reducing and managing behaviour that challenges

A.6.1 Parent training versus any control

Table O.21: Parent training versus any control

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With any control	With parent training		Risk with any control	Risk difference with parent training (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
841 (14 studies)	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	undetected	⨁⨁⨁⨀ MODERATE[1] due to risk of bias	390	451	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.41 standard deviations lower (0.58 to 0.24 lower)
Targeted behaviour that challenges (severity) – follow-up (Better indicated by lower values)											
342 (3 studies) 26- 52 weeks	serious[1]	no serious inconsistency	serious[2]	serious[3]	reporting bias strongly suspected[4]	⨁⨀⨀⨀ VERY LOW[1,2,3],4 due to risk of bias, indirectness, imprecision, publication bias	156	186	-		The mean targeted behaviour that challenges (severity) – follow-up in the intervention groups was 0.13 standard deviations lower (0.34 lower to 0.08 higher)

Quality assessment							Summary of findings				
Targeted behaviour that challenges (severity, non-improvement) – post-treatment											
428 (8 studies)	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	undetected	⨁⨁⨁⨀ MODERATE[1] due to risk of bias	174/197 (88.3 %)	131/231 (56.7 %)	RR 0.67 (0.59 to 0.77)	883 per 1000	291 fewer per 1000 (from 203 fewer to 362 fewer)
Targeted behaviour that challenges (frequency) – post-treatment (Better indicated by lower values)											
633 (9 studies)	serious[1]	serious[5]	no serious indirectness	no serious imprecision	undetected	⨁⨁⨀⨀ LOW[1,5] due to risk of bias, inconsistency	294	339	-		The mean targeted behaviour that challenges (frequency) – post-treatment in the intervention groups was 0.54 standard deviations lower (0.8 to 0.28 lower)
Targeted behaviour that challenges (frequency) – follow-up (Better indicated by lower values)											
258 (12 studies) 26 weeks	serious[6]	no serious inconsistency	no serious indirectness	serious[7]	reporting bias strongly suspected[4]	⨁⨀⨀⨀ VERY LOW[4,6,7] due to risk of bias, imprecision, publication bias	123	135	-		The mean targeted behaviour that challenges (frequency) – follow-up in the intervention groups was 0.23 standard deviations lower (0.47 lower to 0.02 higher)
Targeted behaviour that challenges (frequency, non-improvement) – post-treatment											
343 (6 studies)	serious[1]	no serious inconsistency	serious[2]	no serious imprecision	undetected	⨁⨁⨀⨀ LOW[1,2] due to risk of bias, indirectness	147/155 (94.8 %)	105/188 (55.9 %)	RR 0.63 (0.55 to 0.73)	948 per 1000	351 fewer per 1000 (from 256 fewer to 427 fewer)
Adaptive functioning (communication) – post-treatment (Better indicated by higher values)											
124 (1 study)	serious[6]	no serious inconsisten	serious[2]	very serious[7]	undetected	⨁⨀⨀⨀ VERY LOW[2,6,7]	49	75	-		The mean adaptive functioning

Quality assessment							Summary of findings				
		cy				due to risk of bias, indirectness, imprecision					(communication) – post-treatment in the intervention groups was 0.47 standard deviations higher (0.11 to 0.84 higher)
Adaptive functioning (total) – post-treatment (Better indicated by higher values)											
135 (2 studies)	serious[1]	no serious inconsistency	serious[2]	serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	53	82	-		The mean adaptive functioning (total) – post-treatment in the intervention groups was 0.51 standard deviations higher (0.15 to 0.86 higher)

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

A.6.2 Individual parent training versus group parent training

Table O.22: Individual parent training versus group parent training

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With group parent training	With individual parent training		Risk with group parent training	Risk difference with individual parent training (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
38 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	15	23	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.38 standard deviations lower (1.04 lower to 0.28 higher)
Targeted behaviour that challenges (severity) – follow-up (Better indicated by lower values)											
38 (1 study) 26 weeks	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	15	23	-		The mean targeted behaviour that challenges (severity) – follow-up in the intervention groups was 0.05 standard deviations lower (0.7 lower to 0.61 higher)
Targeted behaviour that challenges (frequency) – post-treatment (Better indicated by lower values)											
31 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk	13	18	-		The mean targeted behaviour that challenges (frequency) – post-treatment in the intervention groups

Quality assessment							Summary of findings				
						of bias, imprecision					was 0.34 standard deviations lower (1.06 lower to 0.38 higher)
Targeted behaviour that challenges (frequency) – follow-up (Better indicated by lower values)											
31 (1 study) 26 weeks	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	13	18	-		The mean targeted behaviour that challenges (frequency) – follow-up in the intervention groups was 0.12 standard deviations higher (0.59 lower to 0.84 higher)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Optimal information size not met; small, single study

A.6.3 Parent training plus optimism training versus parent training alone

Table O.23: Parent training plus optimism training versus parent training alone

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With parent training alone	With parent training plus optimism training		Risk with parent training alone	Risk difference with parent training plus optimism training (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
35 (1 study)	very serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk	17	18	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups

Quality assessment							Summary of findings				
						of bias, imprecision					was 0.8 standard deviations lower (1.49 to 0.11 lower)
Targeted behaviour that challenges (severity, non-improvement) – post-treatment											
35 (1 study)	very serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	11/17 (64.7%)	5/18 (27.8%)	RR 0.43 (0.19 to 0.98)	647 per 1000	369 fewer per 1000 (from 13 fewer to 524 fewer)
Carer satisfaction – post-treatment (Better indicated by higher values)											
35 (1 study)	very serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	17	18	-		The mean carer satisfaction – post-treatment in the intervention groups was 0.22 standard deviations higher (0.44 lower to 0.89 higher)

[1] Crucial limitation for one or more criteria sufficient to substantially lower ones confidence in the estimate of effect
[2] Optimal information size not met; small, single study

A.6.4 Enhanced parent training versus standard parent training

Table O.24: Enhanced parent training versus standard parent training

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With standard parent training	With enhanced parent training		Risk with standard parent training	Risk difference with enhanced parent training (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
50 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecision	26	24	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.06 standard deviations lower (0.62 lower to 0.49 higher)
Targeted behaviour that challenges (severity) – follow-up (Better indicated by lower values)											
42 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecision	19	23	-		The mean targeted behaviour that challenges (severity) – follow-up in the intervention groups was 0.56 standard deviations lower (1.18 lower to 0.06 higher)
Targeted behaviour that challenges (severity, non-improvement) – post-treatment											
50 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecisio	10/26 (38.5%)	13/24 (54.2%)	RR 1.41 (0.77 to	385 per 1000	158 more per 1000 (from 88 fewer to 612 more)

Quality assessment							Summary of findings				
						n			2.59)		
Targeted behaviour that challenges (severity, non-improvement) – follow-up											
42 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecision	11/19 (57.9%)	12/23 (52.2%)	RR 0.9 (0.52 to 1.56)	579 per 1000	58 fewer per 1000 (from 278 fewer to 324 more)
Targeted behaviour that challenges (frequency) – post-treatment (Better indicated by lower values)											
50 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecision	26	24	-		The mean targeted behaviour that challenges (frequency) – post-treatment in the intervention groups was 0.04 standard deviations higher (0.52 lower to 0.59 higher)
Targeted behaviour that challenges (frequency) – follow-up (Better indicated by lower values)											
42 (1 study) 52 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecision	19	23	-		The mean targeted behaviour that challenges (frequency) – follow-up in the intervention groups was 0.04 standard deviations higher (0.56 lower to 0.65 higher)
Targeted behaviour that challenges (frequency, non-improvement) – post-treatment											
50 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecision	11/26 (42.3%)	8/24 (33.3%)	RR 0.79 (0.38 to 1.62)	423 per 1000	89 fewer per 1000 (from 262 fewer to 262 more)
Targeted behaviour that challenges (frequency, non-improvement) – follow-up											
42	no	no serious	no serious	very	undetect	⨁⨁⨀⨀	4/19	8/23	RR	211 per	137 more per 1000

Quality assessment							Summary of findings				
(1 study) 52 weeks	serious risk of bias	inconsistency	indirectness	serious[1]	ed	LOW[1] due to imprecision	(21.1%)	(34.8%)	1.65 (0.59 to 4.65)	1000	(from 86 fewer to 768 more)
Carer satisfaction- post-treatment (Better indicated by higher values)											
50 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecision	26	24	-		The mean carer satisfaction- post-treatment in the intervention groups was 0.18 standard deviations higher (0.38 lower to 0.74 higher)

[1] Optimal information size not met; small, single study

A.7 Psychosocial interventions aimed at reducing and managing behaviour that challenges

A.7.1 Cognitive behavioural interventions versus any control

Table O.25: Cognitive behaviour interventions versus any control

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With any control	With cognitive behavioural interventions		Risk with any control	Risk difference with cognitive behavioural interventions (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (measured with: Family carer rated; Better indicated by lower values)											
103	no serious	no serious inconsistenc	no serious indirectnes	very	undetect	⨁⨁⨀⨀ LOW[1]	58	45	-		The mean targeted behaviour that challenges

Quality assessment							Summary of findings				
(1 study)	risk of bias	y	s	serious[1]	ed	due to imprecision					(severity) – post-treatment in the intervention groups was 0.24 standard deviations lower (0.63 lower to 0.15 higher)
Targeted behaviour that challenges (severity) – follow-up (measured with: Family carer rated; Better indicated by lower values)											
83 (1 study) 31 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨁⨀⨀ LOW[1] due to imprecision	41	42	-		The mean targeted behaviour that challenges (severity) – follow-up in the intervention groups was 0.03 standard deviations lower (0.46 lower to 0.4 higher)
Targeted behaviour that challenges (severity, non-improvement) – post-treatment (assessed with: Paid carer rated)											
38 (1 study)	serious[2]	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⨁⨀⨀⨀ VERY LOW[1,2] due to risk of bias, imprecision	15/20 (75%)	9/18 (50%)	RR 0.67 (0.39 to 1.13)	750 per 1000	247 fewer per 1000 (from 458 fewer to 97 more)
Targeted behaviour that challenges (severity) – post-treatment (measured with: Paid carer rated; Better indicated by lower values)											
194 (2 studies)	no serious risk of bias	serious[3]	no serious indirectness	serious[4]	undetected	⨁⨁⨀⨀ LOW[3,4] due to inconsistency, imprecision	102	92	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.03 standard deviations lower (0.48 lower to 0.42 higher)
Targeted behaviour that challenges (severity) – follow-up (measured with: Paid carer rated; Better indicated by lower values)											
176	no	serious[3]	no serious	serious[4]	undetect	⨁⨁⨀⨀	86	90	-		The mean targeted

Quality assessment							Summary of findings				
(2 studies) 17- 31 weeks	serious risk of bias		indirectness		ed	LOW[3,4] due to inconsistency, imprecision					behaviour that challenges (severity) – follow-up in the intervention groups was 0.13 standard deviations lower (0.58 lower to 0.33 higher)
Adaptive functioning – post-treatment (measured with: Paid carer rated; Better indicated by higher values)											
28 (1 study)	serious [2]	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⊕⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	10	18	-		The mean adaptive functioning – post-treatment in the intervention groups was 1.32 standard deviations higher (0.46 to 2.18 higher)
Quality of life – post-treatment (measured with: Self rated; Better indicated by higher values)											
129 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⊕⊕⊖⊖ LOW[1] due to imprecision	67	62	-		The mean quality of life – post-treatment in the intervention groups was 0.16 standard deviations lower (0.5 lower to 0.19 higher)
Quality of life – follow-up (measured with: Self rated; Better indicated by lower values)											
140 (1 study) 31 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	undetected	⊕⊕⊖⊖ LOW[1] due to imprecision	70	70	-		The mean quality of life – follow-up in the intervention groups was 0.02 standard deviations lower (0.35 lower to 0.32 higher)

[1] Optimal information size not met; small, single study
[2] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[3] $I^2 > 40\%$
[4] Optimal information size not met

A.7.2 Behavioural therapy versus any control

Table O.26: Behavioural therapy versus any control

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With Any control	With Behavioural therapy		Risk with Any control	Risk difference with Behavioural therapy (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
61 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	30	31	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.47 standard deviations lower (0.98 lower to 0.04 higher)
Targeted behaviour that challenges (severity) – follow-up (Better indicated by lower values)											
63 (1 study) 78 weeks	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	30	33	-		The mean targeted behaviour that challenges (severity) – follow-up in the intervention groups was 0.33 standard deviations lower (0.85 lower to 0.19 higher)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Optimal information size not met; small, single study

A.8 Sleep interventions aimed at reducing and managing behaviour that challenges

A.8.1 Sleep interventions versus any control

Table O.27: Sleep interventions versus any control

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With any control	With sleep interventions		Risk with any control	Risk difference with sleep interventions (95% CI)
Targeted behaviour that challenges (global problem sleep behaviour, non-improvement) – post-treatment											
69 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	21/34 (61.8 %)	5/35 (14.3%)	RR 0.23 (0.1 to 0.54)	618 per 1000	476 fewer per 1000 (from 284 fewer to 556 fewer)
Targeted behaviour that challenges (global problem sleep behaviour) – post-treatment (Better indicated by lower values)											
154 (4 studies)	serious[4]	no serious inconsistency	no serious indirectness	serious[5]	undetected	⨁⨁⊖⊖ LOW[4,5] due to risk of bias, imprecision	77	77	-		The mean targeted behaviour that challenges (global problem sleep behaviour) – post-treatment in the intervention groups was 1.05 standard deviations lower (1.48 to 0.63 lower)
Targeted behaviour that challenges (global problem sleep behaviour) – follow-up (Better indicated by lower values)											
130	serio	serious[6]	no serious	serious[5]	undetect	⨁⊖⊖⊖	55	75	-		The mean targeted behaviour

Quality assessment							Summary of findings				
(3 studies) 6 to 26 weeks	us[4]		indirectness		ed	VERY LOW[4,5,6] due to risk of bias, inconsistency, imprecision					that challenges (global problem sleep behaviour) – follow-up in the intervention groups was 0.92 standard deviations lower (1.6 to 0.24 lower)
Targeted behaviour that challenges (total sleep time) – post-treatment (measured with: Actigraph; Better indicated by higher values)											
96 (2 studies)	serious[4]	no serious inconsistency	no serious indirectness	serious[5]	undetected	⊕⊕⊖⊖ LOW[4,5] due to risk of bias, imprecision	48	48	-		The mean targeted behaviour that challenges (total sleep time) – post-treatment in the intervention groups was 0.62 standard deviations higher (0.2 to 1.03 higher)
Targeted behaviour that challenges (sleep efficiency) – post-treatment (measured with: Actigraph; Better indicated by higher values)											
96 (2 studies)	serious[4]	no serious inconsistency	no serious indirectness	serious[5]	undetected	⊕⊕⊖⊖ LOW[4,5] due to risk of bias, imprecision	48	48	-		The mean targeted behaviour that challenges (sleep efficiency) – post-treatment in the intervention groups was 0.24 standard deviations higher (0.26 lower to 0.74 higher)
Targeted behaviour that challenges (total sleep time) – follow-up (measured with: Actigraph; Better indicated by higher values)											
46 (1 study) 26 weeks	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⊕⊖⊖⊖ VERY LOW[1,3] due to risk of bias, imprecision	23	23	-		The mean targeted behaviour that challenges (total sleep time) – follow-up in the intervention groups was 0.14 standard deviations higher (0.44 lower to 0.71 higher)

Quality assessment							Summary of findings				
Targeted behaviour that challenges (sleep efficiency) – follow-up (measured with: Actigraph; Better indicated by lower values)											
46 (1 study) 26 weeks	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⊕⊖⊖⊖ VERY LOW[1,3] due to risk of bias, imprecision	23	23	-		The mean targeted behaviour that challenges (sleep efficiency) – follow-up in the intervention groups was 0.11 standard deviations lower (0.69 lower to 0.46 higher)
Targeted behaviour that challenges (sleep onset latency) – post-treatment (measured with: Actigraph; Better indicated by lower values)											
69 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⊕⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	34	35	-		The mean targeted behaviour that challenges (sleep onset latency) – post-treatment in the intervention groups was 0.59 standard deviations lower (1.07 to 0.11 lower)
Targeted behaviour that challenges (wake after sleep onset) – post-treatment (measured with: Actigraph; Better indicated by lower values)											
96 (2 studies)	serious[4]	serious[6]	no serious indirectness	serious[5]	undetected	⊕⊖⊖⊖ VERY LOW[4,5,6] due to risk of bias, inconsistency, imprecision	48	48	-		The mean targeted behaviour that challenges (wake after sleep onset) – post-treatment in the intervention groups was 0.31 standard deviations lower (1.13 lower to 0.51 higher)
Targeted behaviour that challenges (wake after sleep onset) – follow-up (measured with: Actigraph; Better indicated by lower values)											
46 (1 study) 26 weeks	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⊕⊖⊖⊖ VERY LOW[1,3] due to risk of bias, imprecision	23	23	-		The mean targeted behaviour that challenges (wake after sleep onset) – follow-up in the intervention groups was 0.29 standard deviations higher (0.29 lower to 0.88 higher)

Quality assessment							Summary of findings				
Targeted behaviour that challenges (total sleep time) post-treatment (measured with: Sleep diary; Better indicated by higher values)											
30 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,3] due to risk of bias, imprecision	15	15	-		The mean targeted behaviour that challenges (total sleep time) post-treatment in the intervention groups was 0.3 standard deviations lower (1.02 lower to 0.42 higher)
Targeted behaviour that challenges (activity score) – post-treatment (measured with: Sleep diary; Better indicated by lower values)											
30 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,3] due to risk of bias, imprecision	15	15	-		The mean targeted behaviour that challenges (activity score) – post-treatment in the intervention groups was 0.28 standard deviations higher (0.44 lower to 1 higher)
Carer Satisfaction (non-satisfied) – post-treatment											
30 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,3] due to risk of bias, imprecision	2/17 (11.8 %)	1/13 (7.7%)	RR 0.65 (0.07 to 6.45)	118 per 1000	41 fewer per 1000 (from 109 fewer to 641 more)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Applicability- different populations
[3] Optimal information size not met; small, single study
[4] Most information is from studies at moderate risk of bias
[5] Optimal information size not met
[6] I^2 > 40%

A.8.2 Face-to-face sleep intervention versus booklet only

Table O.28: Face-to-face sleep intervention versus booklet only

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With booklet only	With face-to-face sleep intervention		Risk with booklet only	Risk difference with face-to-face sleep intervention (95% CI)
Targeted behaviour that challenges (global problem sleep behaviour) – follow-up (Better indicated by lower values)											
42 (1 study) 26 weeks	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	22	20	-		The mean targeted behaviour that challenges (global problem sleep behaviour) – follow-up in the intervention groups was 0.07 standard deviations lower (0.68 lower to 0.53 higher)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Optimal information size not met; small, single study

A.9 Pharmacological interventions aimed at reducing and managing behaviour that challenges

A.9.1 Risperidone versus placebo in children and young people

Table O.29: Risperidone versus placebo in children and young people

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With placebo	With risperidone		Risk with placebo	Risk difference with risperidone (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (measured with: End-point score; Better indicated by lower values)											
257 (4 studies)	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	undetected	⨁⨁⨀⨀ LOW[1,2] due to risk of bias, imprecision	141	116	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 1.09 standard deviations lower (1.39 to 0.79 lower)
Targeted behaviour that challenges (severity) – post-treatment (measured with: Change score; Better indicated by lower values)											
66 (1 study)	serious[3]	no serious inconsistency	serious[4]	very serious[5]	undetected	⨁⨀⨀⨀ VERY LOW[3,4,5] due to risk of bias, indirectness, imprecision	35	31	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.98 standard deviations lower (1.49 to 0.47 lower)

Quality assessment							Summary of findings				
Targeted behaviour that challenges (severity, non-improvement) – post-treatment											
153 (2 studies)	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	undetected	⨁⨁⨀⨀ LOW[1,2] due to risk of bias, imprecision	68/80 (85%)	25/73 (34.2%)	RR 0.42 (0.28 to 0.64)	850 per 1000	493 fewer per 1000 (from 306 fewer to 612 fewer)
Adaptive functioning (social) – post-treatment (measured with: Nisonger Child Behaviour Rating Form – Social Compliance[6]; Better indicated by higher values)											
155 (3 studies)	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	undetected	⨁⨁⨀⨀ LOW[1,2] due to risk of bias, imprecision	88	67	-		The mean adaptive functioning (social) – post-treatment in the intervention groups was 0.86 standard deviations higher (0.42 to 1.3 higher)
Adverse events (elevated prolactin, non-occurrence) – post-treatment											
228 (2 studies)	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	undetected	⨁⨁⨀⨀ LOW[1,2] due to risk of bias, imprecision	119/120 (99.2%)	97/108 (89.8%)	RR 0.91 (0.85 to 0.97)	992 per 1000	89 fewer per 1000 (from 30 fewer to 149 fewer)
Adverse events (prolactin-related adverse event; oligomenorrhea, non-occurrence) – post-treatment											
66 (1 study)	serious[3]	no serious inconsistency	serious[4]	very serious[5]	undetected	⨁⨀⨀⨀ VERY LOW[3,4,5] due to risk of bias, indirectness, imprecision	35/35 (100%)	30/31 (96.8%)	RR 0.97 (0.89 to 1.05)	1000 per 1000	30 fewer per 1000 (from 110 fewer to 50 more)
Adverse events (prolactin level; ng/ml) – post-treatment (Better indicated by lower values)											
241 (3 studies)	serious[3]	no serious inconsistency	serious[4]	serious[2]	undetected	⨁⨀⨀⨀ VERY LOW[2,3,4] due to risk of	125	116	-		The mean adverse events (prolactin level; ng/ml) – post-treatment in the

Quality assessment							Summary of findings				
						bias, indirectness, imprecision					intervention groups was 3.22 standard deviations higher (1.68 to 4.75 higher)
Adverse events (weight; kg) – post-treatment (measured with: Change score; Better indicated by lower values)											
282 (3 studies)	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	undetected	⨁⨁⊖⊖ LOW[1,2] due to risk of bias, imprecision	150	132	-		The mean adverse events (weight; kg) – post-treatment in the intervention groups was 0.82 standard deviations higher (0.57 to 1.06 higher)
Adverse events (weight; kg) – post-treatment (measured with: Endpoint score; Better indicated by lower values)											
53 (1 study)	serious[3]	no serious inconsistency	serious[4]	very serious[5]	undetected	⨁⊖⊖⊖ VERY LOW[3,4,5] due to risk of bias, indirectness, imprecision	28	25	-		The mean adverse events (weight; kg) – post-treatment in the intervention groups was 0.39 standard deviations higher (0.16 lower to 0.93 higher)
Adverse events (weight gain, non-occurrence) – post-treatment											
277 (3 studies)	serious[1]	no serious inconsistency	serious[4]	serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,4] due to risk of bias, indirectness, imprecision	147/148 (99.3%)	115/129 (89.1%)	RR 0.91 (0.85 to 0.96)	993 per 1000	89 fewer per 1000 (from 40 fewer to 149 fewer)
Adverse events (somnolence/sedation, non-occurrence) – post-treatment											
550 (6 studies)	serious[1]	serious[7]	serious[4]	no serious imprecision	undetected	⨁⊖⊖⊖ VERY LOW[1,4,7] due to risk of bias,	249/283 (88%)	138/267 (51.7%)	RR 0.58 (0.44 to	880 per 1000	370 fewer per 1000 (from 202 fewer to 493 fewer)

Quality assessment							Summary of findings				
						inconsistency, indirectness			0.77)		
Adverse events (seizure, non-occurrence) – post-treatment											
101 (1 study)	serious[3]	no serious inconsistency	no serious indirectness	very serious[5]	undetected	⨁⊖⊖⊖ VERY LOW[3,5] due to risk of bias, imprecision	51/52 (98.1 %)	49/49 (100%)	RR 1.02 (0.97 to 1.08)	981 per 1000	20 more per 1000 (from 29 fewer to 78 more)
Adverse events (discontinuation due to adverse events, non-occurrence) – post-treatment											
340 (4 studies)	serious[1]	no serious inconsistency	serious[4]	no serious imprecision[2]	undetected	⨁⨁⊖⊖ LOW[1,2,4] due to risk of bias, indirectness	175/178 (98.3 %)	158/162 (97.5%)	RR 0.99 (0.96 to 1.03)	983 per 1000	10 fewer per 1000 (from 39 fewer to 29 more)
Adverse events (discontinuation due other reasons, non-occurrence) – post-treatment											
450 (5 studies)	serious[1]	serious[7]	serious[4]	no serious imprecision	undetected	⨁⊖⊖⊖ VERY LOW[1,4,7] due to risk of bias, inconsistency, indirectness	170/235 (72.3 %)	190/215 (88.4%)	RR 1.19 (1.06 to 1.34)	723 per 1000	137 more per 1000 (from 43 more to 246 more)

1 Most information is from studies at moderate risk of bias
[2] Optimal information size not met
[3] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[4] Applicability – different populations
[5] Optimal information size not met; small, single study
[6] Combined adaptive social and compliant/calm subscales
[7] I^2 > 40%

A.9.2 Withdrawal of risperidone versus continuation of risperidone in children and young people

Table O.30: Withdrawal of risperidone versus continuation of risperidone in children and young people

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With continuation of risperidone	With withdrawal of risperidone		Risk with continuation of risperidone	Risk difference with withdrawal of risperidone (95% CI)
Targeted behaviour that challenges (relapse) – post-treatment											
32 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	2/16 (12.5%)	10/16 (62.5%)	RR 5 (1.3 to 19.3)	125 per 1000	500 more per 1000 (from 37 more to 1000 more)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Applicability – different populations
[3] Optimal information size not met; small, single study

A.9.3 Aripiprazole versus placebo in children and young people

Table O.31: Aripiprazole versus placebo in children and young people

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With placebo	With aripiprazole		Risk with placebo	Risk difference with aripiprazole (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
308 (2 studies)	serious[1]	no serious inconsistency	serious[2]	serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	98	210	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.64 standard deviations lower (0.91 to 0.36 lower)
Targeted behaviour that challenges (severity, non-improvement) – post-treatment											
308 (2 studies)	serious[1]	no serious inconsistency	serious[2]	serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	74/98 (75.5 %)	100/210 (47.6%)	RR 0.65 (0.5 to 0.84)	755 per 1000	264 fewer per 1000 (from 121 fewer to 378 fewer)
Quality of life – post-treatment (Better indicated by higher values)											
243 (2 studies)	serious[1]	very serious[4]	serious[2]	serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3],4 due to risk of bias, inconsistency, indirectness, imprecision	76	167	-		The mean quality of life – post-treatment in the intervention groups was 0.6 standard deviations higher (0.17 lower to 1.37 higher)

Quality assessment							Summary of findings				
Adverse events (elevated prolactin, non-occurrence) – post-treatment											
313 (2 studies)	serious[1]	no serious inconsistency	serious[2]	serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	96/101 (95%)	211/212 (99.5%)	RR 1.05 (0.99 to 1.1)	950 per 1000	48 more per 1000 (from 10 fewer to 95 more)
Adverse events (weight gain; kg) – post- treatment (Better indicated by lower values)											
216 (1 study)	serious[5]	no serious inconsistency	serious[2]	very serious[6]	undetected	⨁⊖⊖⊖ VERY LOW[2,5,6] due to risk of bias, indirectness, imprecision	51	165	-		The mean adverse events (weight gain; kg) – post-treatment in the intervention groups was 0.48 standard deviations higher (0.17 to 0.8 higher)
Adverse events (weight gain; clinically sig., non-occurrence)											
313 (2 studies)	serious[1]	no serious inconsistency	serious[2]	serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	94/101 (93.1%)	156/212 (73.6%)	RR 0.79 (0.71 to 0.88)	931 per 1000	195 fewer per 1000 (from 112 fewer to 270 fewer)
Adverse events (sedation, non-occurrence) – post-treatment											
313 (2 studies)	serious[1]	no serious inconsistency	serious[2]	serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	96/101 (95%)	165/212 (77.8%)	RR 0.83 (0.76 to 0.91)	950 per 1000	162 fewer per 1000 (from 86 fewer to 228 fewer)
Adverse events (seizure, non-occurrence) – post-treatment											
216 (1 study)	serious[5]	no serious inconsistency	serious[2]	very serious[6]	undetected	⨁⊖⊖⊖ VERY LOW[2,5,6] due to risk of bias, indirectness, imprecision	50/51 (98%)	165/165 (100%)	RR 1.03 (0.98 to 1.08)	980 per 1000	29 more per 1000 (from 20 fewer to 78 more)

Quality assessment							Summary of findings				
Adverse events (discontinuation due to adverse events, non-occurrence) – post-treatment											
316 (2 studies)	serious[1]	no serious inconsistency	serious[2]	serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	96/103 (93.2%)	191/213 (89.7%)	RR 0.96 (0.89 to 1.04)	932 per 1000	37 fewer per 1000 (from 103 fewer to 37 more)
Adverse events (discontinuation due to other reasons, non-occurrence) – post-treatment											
316 (2 studies)	serious[1]	no serious inconsistency	serious[2]	serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	81/103 (78.6%)	201/213 (94.4%)	RR 1.19 (1.07 to 1.33)	786 per 1000	149 more per 1000 (from 55 more to 260 more)

1 Most information is from studies at moderate risk of bias
[2] Applicability – different populations
[3] Optimal information size not met
[4] I^2 > 75%
[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

A.9.4 Aripiprazole versus risperidone in children and young people

Table O.32: Aripiprazole versus risperidone in children and young people

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With Risperidone	With Aripiprazole		Risk with Risperidone	Risk difference with Aripiprazole (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
59 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	30	29	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.38 standard deviations higher (0.14 lower to 0.9 higher)
Adverse events (drowsiness, non-occurrence) – post-treatment											
59 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	25/30 (83.3%)	23/29 (79.3%)	RR 0.95 (0.74 to 1.22)	833 per 1000	42 fewer per 1000 (from 217 fewer to 183 more)
Adverse events (seizure, non-occurrence) – post-treatment											
59 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	29/30 (96.7%)	29/29 (100%)	RR 1.03 (0.94 to 1.13)	967 per 1000	29 more per 1000 (from 58 fewer to 126 more)

Quality assessment							Summary of findings				
Adverse events (discontinuation due to adverse events, non-occurrence) – post-treatment											
59 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	29/30 (96.7%)	29/29 (100%)	RR 1.03 (0.94 to 1.13)	967 per 1000	29 more per 1000 (from 58 fewer to 126 more)
Adverse events (discontinuation due to other reasons, non-occurrence) – post-treatment											
59 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	28/30 (93.3%)	27/29 (93.1%)	RR 1 (0.87 to 1.14)	933 per 1000	0 fewer per 1000 (from 121 fewer to 131 more)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Applicability – different populations
[3] Optimal information size not met; small, single study

A.9.5 Withdrawal of aripiprazole versus continuation of aripiprazole in children and young people

Table O.33: Withdrawal of aripiprazole versus continuation of aripiprazole in children and young people

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With continuation of aripiprazole	With withdrawal of aripiprazole		Risk with continuation of aripiprazole	Risk difference with withdrawal of aripiprazole (95% CI)
Targeted behaviour that challenges (relapse) – post-treatment											
85 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	14/41 (34.1%)	23/44 (52.3%)	RR 1.53 (0.92 to 2.55)	341 per 1000	181 more per 1000 (from 27 fewer to 529 more)
Adverse events (weight gain; clinically sig., non-occurrence)											
85 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	39/41 (95.1%)	43/44 (97.7%)	RR 1.03 (0.95 to 1.12)	951 per 1000	29 more per 1000 (from 48 fewer to 114 more)
Adverse events (discontinuation due to adverse events, non-occurrence) – post-treatment											
85 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	41/41 (100%)	43/44 (97.7%)	RR 0.98 (0.92 to 1.04)	1000 per 1000	20 fewer per 1000 (from 80 fewer to 40 more)

Quality assessment							Summary of findings				
Adverse events (discontinuation due to other reasons, non-occurrence) – post-treatment											
85 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	22/41 (53.7%)	20/44 (45.5%)	RR 0.85 (0.55 to 1.3)	537 per 1000	80 fewer per 1000 (from 241 fewer to 161 more)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Applicability – different populations
[3] Optimal information size not met; small, single study

A.9.6 Olanzapine versus haloperidol in children and young people

Table O.34: Olanzapine versus haloperidol in children and young people

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With haloperidol	With olanzapine		Risk with haloperidol	Risk difference with olanzapine (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
12 (1 study)	very serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	6	6	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 1.4 standard deviations lower (2.73 to 0.08 lower)
Adverse events (drowsiness, non-occurrence) – post-treatment											
12	very	no serious	no serious	very	undetect	⨁⊖⊖⊖	4/6	1/6	RR	667 per	500 fewer per 1000

(1 study)	serious[1]	inconsistency	indirectness	serious[2]	ed	VERY LOW[1,2] due to risk of bias, imprecision	(66.7%)	(16.7%)	0.25 (0.04 to 1.63)	1000	(from 640 fewer to 420 more)
Adverse events – (weight gain; kg) – post-treatment (Better indicated by lower values)											
12 (1 study)	very serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	6	6	-		The mean adverse events – (weight gain; kg) – post-treatment in the intervention groups was 1.26 standard deviations higher (0.03 lower to 2.54 higher)
Adverse events (weight gain) – post-treatment											
12 (1 study)	very serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	6/6 (100%)	5/6 (83.3%)	RR 0.85 (0.55 to 1.31)	1000 per 1000	150 fewer per 1000 (from 450 fewer to 310 more)

[1] Crucial limitation for one or more criteria sufficient to substantially lower ones confidence in the estimate of effect.
[2] Optimal information size not met; small, single study

A.9.7 Topiramate (plus risperidone) versus placebo (plus risperidone) in children and young people

Table O.35: Topiramate (plus risperidone) versus placebo (plus risperidone) in children and young people

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With placebo plus risperidone	With topiramate plus risperidone		Risk with Placebo plus risperidone	Risk difference with topiramate plus risperidone (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
40 (1 study)	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to indirectness, imprecision	20	20	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 1.88 standard deviations lower (2.63 to 1.12 lower)
Adverse events (sedation, non-occurrence) – post-treatment											
40 (1 study)	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to indirectness, imprecision	16/20 (80%)	19/20 (95%)	RR 1.19 (0.93 to 1.51)	800 per 1000	152 more per 1000 (from 56 fewer to 408 more)
Adverse events (weight at endpoint; kg) – post-treatment (Better indicated by lower values)											
40 (1 study)	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to	20	20	-		The mean adverse events (weight at endpoint; kg) – post-treatment in the

Quality assessment							Summary of findings				
						indirectness, imprecision					intervention groups was 0.24 standard deviations lower (0.87 lower to 0.38 higher)

[1] Applicability – different populations
[2] Optimal information size not met; small, single study

A.9.8 Valproate versus placebo in children and young people

Table O.36: Topiramate (plus risperidone) versus placebo (plus risperidone) in children and young people

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With placebo	With valproate		Risk with placebo	Risk difference with valproate (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
57 (2 studies)	serious[1]	serious[2]	no serious indirectness	serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, inconsistency, imprecision	25	32	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.06 standard deviations lower (0.75 lower to 0.63 higher)
Targeted behaviour that challenges (severity, non-improvement) – post-treatment											
27 (1 study)	serious[4]	no serious inconsistency	no serious indirectness	very serious[5]	undetected	⨁⊖⊖⊖ VERY LOW[4,5] due to risk of bias, imprecision	10/11 (90.9 %)	6/16 (37.5 %)	RR 0.41 (0.21 to 0.8)	909 per 1000	536 fewer per 1000 (from 182 fewer to 718 fewer)

Adverse events (weight gain; kg) – post-treatment (measured with: Change score; Better indicated by lower values)											
57 (2 studies)	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	undetected	⨁⨁⨀⨀ LOW[1,3] due to risk of bias, imprecision	25	32	-		The mean adverse events (weight gain; kg) – post-treatment in the intervention groups was 0.29 standard deviations higher (0.24 lower to 0.82 higher)
Adverse events (weight gain, non-occurrence) – post-treatment											
30 (1 study)	serious[4]	no serious inconsistency	no serious indirectness	very serious[5]	undetected	⨁⨀⨀⨀ VERY LOW[4,5] due to risk of bias, imprecision	10/14 (71.4 %)	9/16 (56.3 %)	RR 0.79 (0.46 to 1.36)	714 per 1000	150 fewer per 1000 (from 386 fewer to 257 more)
Adverse events (somnolence/sedation, non-occurrence) – post-treatment											
57 (2 studies)	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	undetected	⨁⨁⨀⨀ LOW[1,3] due to risk of bias, imprecision	19/25 (76%)	29/32 (90.6 %)	RR 1.19 (0.9 to 1.56)	760 per 1000	144 more per 1000 (from 76 fewer to 426 more)
Adverse events (discontinuation due to adverse events, non-occurrence) – post-treatment											
57 (2 studies)	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	undetected	⨁⨁⨀⨀ LOW[1,3] due to risk of bias, imprecision	25/25 (100 %)	30/32 (93.8 %)	RR 0.95 (0.83 to 1.08)	1000 per 1000	50 fewer per 1000 (from 170 fewer to 80 more)
Adverse events (discontinuation due to other reasons, non-occurrence) – post-treatment											
27 (1 study)	serious[4]	no serious inconsistency	no serious indirectness	very serious[5]	undetected	⨁⨀⨀⨀ VERY LOW[4,5] due to risk of bias, imprecision	10/11 (90.9 %)	15/16 (93.8 %)	RR 1.03 (0.82 to 1.29)	909 per 1000	27 more per 1000 (from 164 fewer to 264 more)

1 Most information is from studies at moderate risk of bias
[2] I^2 > 40%

[3] Optimal information size not met
[4] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[5] Optimal information size not met; small, single study

A.9.9 N-acetylcysteine versus placebo in children and young people

Table O.37: N-acetylcysteine versus placebo in children and young people

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With placebo	With N-acetylcysteine (NAC)		Risk with placebo	Risk difference with N-acetylcysteine (NAC) (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
29 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	15	14	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.70 standard deviations lower (1.46 lower to 0.05 higher)
Adverse events (discontinuation due to adverse events, non-occurrence) – post-treatment											
33 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	18/18 (100 %)	14/15 (93.3%)	RR 0.93 (0.78 to 1.11)	1000 per 1000	70 fewer per 1000 (from 220 fewer to 110 more)
Adverse events (discontinuation due to other reasons, non-occurrence) – post-treatment											
33	serio	no serious	serious[2]	very	undetect	⨁⊖⊖⊖	12/18	14/15	RR 1.4	667	267 more per 1000

Quality assessment							Summary of findings				
(1 study)	us[1]	inconsistency		serious[3]	ed	VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	(66.7 %)	(93.3%)	(0.98 to 1.99)	per 1000	(from 13 fewer to 660 more)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Applicability – different populations
[3] Optimal information size not met; small, single study

A.9.10 Ginkgo biloba (plus risperidone) versus placebo (plus risperidone) in children and young people

Table O.38: Ginkgo biloba (plus risperidone) versus placebo (plus risperidone) in children and young people

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With placebo plus risperidone	With ginkgo biloba plus risperidone		Risk with placebo plus risperidone	Risk difference with ginkgo biloba plus risperidone (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
47 (1 study)	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to indirectness, imprecision	24	23	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.1 standard deviations higher (0.47 lower to 0.67 higher)

Quality assessment							Summary of findings				
Adverse events (drowsiness, non-occurrence) – post-treatment											
47 (1 study)	no serious risk of bias	no serious inconsistency	serious[1]	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to indirectness, imprecision	17/24 (70.8%)	17/23 (73.9%)	RR 1.04 (0.73 to 1.49)	708 per 1000	28 more per 1000 (from 191 fewer to 347 more)

[1] Applicability – different populations
[2] Optimal information size not met; small, single study

A.9.11 Omega-3 versus placebo in children and young people

Table O.39: Omega-3 versus placebo in children and young people

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With placebo	With omega-3		Risk with placebo	Risk difference with omega-3 (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
12 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	5	7	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.37 standard deviations higher (0.79 lower to 1.53 higher)
Adverse events (discontinuation due to adverse events, non-occurrence) – post-treatment											
13	seriou	no serious	serious[2]	very	undetect	⨁⊖⊖⊖	5/6	7/7	RR	833	158 more per 1000

Quality assessment							Summary of findings				
(1 study)	s[1]	inconsistency		serious[3]	ed	VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	(83.3 %)	(100 %)	1.19 (0.78 to 1.83)	per 1000	(from 183 fewer to 692 more)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Applicability – different populations
[3] Optimal information size not met; small, single study

A.9.12 Piracetam (plus risperidone) versus placebo (plus risperidone) in children and young people

Table O.40: Piracetam (plus risperidone) versus placebo (plus risperidone) in children and young people

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With placebo (plus risperidone)	With piracetam (plus risperidone)		Risk with placebo (plus risperidone)	Risk difference with piracetam (plus risperidone) (95% CI)
Adverse events (drowsiness, non-occurrence) – post-treatment											
40 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	11/20 (55%)	13/20 (65%)	RR 1.18 (0.71 to 1.97)	550 per 1000	99 more per 1000 (from 160 fewer to 534 more)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Applicability – different populations
[3] Optimal information size not met; small, single study

A.9.13 Risperidone versus placebo in adults

Table O.41: Risperidone versus placebo in adults

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With placebo	With risperidone		Risk with placebo	Risk difference with risperidone (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (measured with: End-point score; 12 week; Better indicated by lower values)											
88 (2 studies)	no serious risk of bias	serious[1]	no serious indirectness	serious[2]	undetected	⨁⨁⨀⨀ LOW[1,2] due to inconsistency, imprecision	45	43	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.25 standard deviations lower (0.94 lower to 0.44 higher)
Targeted behaviour that challenges (severity) – post-treatment (measured with: Change-score; 12 week; Better indicated by lower values)											
74 (1 study)	serious[3]	no serious inconsistency	no serious indirectness	very serious[4]	undetected	⨁⨀⨀⨀ VERY LOW[3,4] due to risk of bias, imprecision	37	37	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.44 standard deviations lower (0.9 lower to 0.02 higher)
Targeted behaviour that challenges (severity) – post-treatment (measured with: Endpoint-score; 26 weeks[5]; Better indicated by lower values)											
37 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	undetected	⨁⨁⨀⨀ LOW[4] due to imprecision	20	17	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.16 standard deviations

Quality assessment							Summary of findings				
											higher (0.48 lower to 0.81 higher)
Quality of life – post-treatment (measured with: 12 weeks; Better indicated by higher values)											
58 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	undetected	⊕⊕⊖⊖ LOW[4] due to imprecision	29	29	-		The mean quality of life – post-treatment in the intervention groups was 0.27 standard deviations higher (0.25 lower to 0.79 higher)
Quality of life – post-treatment (measured with: 26 weeks[5]; Better indicated by higher values)											
40 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	undetected	⊕⊕⊖⊖ LOW[4] due to imprecision	21	19	-		The mean quality of life – post-treatment in the intervention groups was 0.2 standard deviations higher (0.42 lower to 0.82 higher)
Adaptive functioning (social) – post-treatment (Better indicated by lower values)											
30 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	undetected	⊕⊕⊖⊖ LOW[4] due to imprecision	16	14	-		The mean adaptive functioning (social) – post-treatment in the intervention groups was 1.36 standard deviations lower (2.17 to 0.56 lower)
Adverse events (weight gain, non-occurrence) – post-treatment											
31 (1 study)	no serious risk of bias	no serious inconsistency	serious[6]	very serious[4]	undetected	⊕⊖⊖⊖ VERY LOW[4,6] due to indirectness, imprecision	16/16 (100 %)	13/15 (86.7%)	RR 0.87 (0.69 to 1.09)	1000 per 1000	130 fewer per 1000 (from 310 fewer to 90 more)
Adverse events (somnolence/sedation, non-occurrence) – post-treatment											
108 (2	no serious	very serious[7]	no serious indirectnes	serious[2]	undetected	⊕⊖⊖⊖ VERY LOW[2,7]	48/54 (88.9	36/54 (66.7%	RR 0.65	889 per	311 fewer per 1000 (from 640 fewer to 418 more)

Quality assessment							Summary of findings				
studies)	risk of bias		s			due to inconsistency, imprecision	%)	)	(0.28 to 1.47)	1000	
Adverse events (discontinuation due to adverse events, non-occurrence) – post-treatment											
89 (2 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[4]	undetected	⨁⨁⨁⊖ MODERATE[4] due to imprecision	45/45 (100 %)	41/44 (93.2%)	RR 0.95 (0.87 to 1.04)	1000 per 1000	50 fewer per 1000 (from 130 fewer to 40 more)
Adverse events (discontinuation due to other reasons, non-occurrence) – post-treatment											
166 (3 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[4]	undetected	⨁⨁⨁⊖ MODERATE[4] due to imprecision	67/83 (80.7 %)	70/83 (84.3%)	RR 1.04 (0.92 to 1.18)	807 per 1000	32 more per 1000 (from 65 fewer to 145 more)

[1] I^2 > 40%
[2] Optimal information size not met
[3] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[4] Optimal information size not met; small, single study
[5] Participants agreed to take the study drug for 12 weeks, with the option of continuing until 26 weeks, unless at 12 weeks other options were preferred. Post-treatment data is therefore provided at both 12 and 26 week end of treatment.
[6] Applicability – different populations
[7] I^2 > 75%

A.9.14 Haloperidol versus placebo in adults

Table O.42: Haloperidol versus placebo in adults

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With placebo	With haloperidol		Risk with placebo	Risk difference with haloperidol (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (measured with: 12 weeks[1]; Better indicated by lower values)											
57 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⨁⊖⊖ LOW[2] due to imprecision	29	28	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.48 standard deviations lower (1 lower to 0.05 higher)
Targeted behaviour that challenges (severity) – post-treatment (measured with: 26 weeks[1]; Better indicated by lower values)											
40 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⨁⊖⊖ LOW[2] due to imprecision	20	20	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.25 standard deviations lower (0.87 lower to 0.37 higher)
Quality of life – post-treatment (measured with: 12 weeks[1]; Better indicated by higher values)											

Quality assessment							Summary of findings				
57 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⨁⨀⨀ LOW[2] due to imprecision	29	28	-		The mean quality of life – post-treatment in the intervention groups was 0.17 standard deviations lower (0.69 lower to 0.35 higher)
Quality of life – post-treatment (measured with: 26 weeks[1]; Better indicated by higher values)											
41 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⨁⨀⨀ LOW[2] due to imprecision	21	20	-		The mean quality of life – post-treatment in the intervention groups was 0.18 standard deviations lower (0.79 lower to 0.43 higher)
Adverse events (seizure, non-occurrence) – post-treatment											
57 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⨁⨀⨀ LOW[2] due to imprecision	29/29 (100%)	27/28 (96.4%)	RR 0.96 (0.88 to 1.06)	1000 per 1000	40 fewer per 1000 (from 120 fewer to 60 more)
Adverse events (discontinuation due to adverse events, non-occurrence) – post-treatment											
57 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⨁⨀⨀ LOW[2] due to imprecision	29/29 (100%)	26/28 (92.9%)	RR 0.93 (0.82 to 1.05)	1000 per 1000	70 fewer per 1000 (from 180 fewer to 50 more)
Adverse events (discontinuation due to other reasons, non-occurrence) – post-treatment											
57	no	no serious	no serious	very	undetecte	⨁⨁⨀⨀	21/29	23/28	RR	724	94 more per 1000

Quality assessment							Summary of findings				
(1 study)	serious risk of bias	inconsistency	indirectness	serious[2]	d	LOW[2] due to imprecision	(72.4%)	(82.1%)	1.13 (0.85 to 1.51)	per 1000	(from 109 fewer to 369 more)

[1] Patients agreed to take the study drug for 12 weeks, with the option of continuing until 26 weeks, unless at 12 weeks other options were preferred. Post-treatment data is therefore provided at both 12 and 26 week end of treatment.

[2] Optimal information size not met; small, single trial

A.9.15 Risperidone versus haloperidol in adults

Table O.43: Risperidone versus haloperidol in adults

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With haloperidol	With risperidone		Risk with haloperidol	Risk difference with risperidone (95% CI)
Targeted behaviour that challenges (severity) – post-treatment (measured with: 12 weeks[1]; Better indicated by lower values)											
57 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⨁⨀⨀ LOW[2] due to imprecision	28	29	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.49 standard deviations higher (0.03 lower to 1.02 higher)
Targeted behaviour that challenges (severity) – post-treatment (measured with: 26 weeks[1]; Better indicated by lower values)											
36 (1 study)	no serious risk of	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⨁⨀⨀ LOW[2] due to	19	17	-		The mean targeted behaviour that challenges (severity) – post-treatment in the

Quality assessment							Summary of findings				
	bias					imprecision					intervention groups was 0.39 standard deviations higher (0.28 lower to 1.05 higher)
Quality of life – post-treatment (measured with: 12 weeks[1]; Better indicated by higher values)											
57 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⨁⨀⨀ LOW[2] due to imprecision	28	29	-		The mean quality of life – post-treatment in the intervention groups was 0.43 standard deviations higher (0.09 lower to 0.96 higher)
Quality of life – post-treatment (measured with: 26 weeks 1; Better indicated by higher values)											
39 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⨁⨀⨀ LOW[2] due to imprecision	20	19	-		The mean quality of life – post-treatment in the intervention groups was 0.41 standard deviations higher (0.23 lower to 1.04 higher)
Adverse events (seizure, non-occurrence) – post-treatment											
57 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⨁⨀⨀ LOW[2] due to imprecision	27/28 (96.4%)	29/29 (100%)	RR 1.04 (0.94 to 1.14)	964 per 1000	39 more per 1000 (from 58 fewer to 135 more)
Adverse events (discontinuation due to adverse events) – post-treatment											
57 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⨁⨀⨀ LOW[2] due to imprecision	26/28 (92.9%)	28/29 (96.6%)	RR 1.04 (0.92 to 1.18)	929 per 1000	37 more per 1000 (from 74 fewer to 167 more)
Adverse events (discontinuation due to other reasons) – post-treatment											
57	no	no serious	no serious	very	undetect	⨁⨁⨀⨀	24/28	23/29	RR	857 per	60 fewer per 1000

Quality assessment							Summary of findings				
(1 study)	serious risk of bias	inconsistency	indirectness	serious[2]	ed	LOW[2] due to imprecision	(85.7%)	(79.3%)	0.93 (0.73 to 1.18)	1000	(from 231 fewer to 154 more)

[1] Patients agreed to take the study drug for 12 weeks, with the option of continuing until 26 weeks, unless at 12 weeks other options were preferred. Post-treatment data is therefore provided at both 12 and 26 week end of treatment.
[2] Optimal information size not met; small, single study

A.9.16 Olanzapine versus risperidone in adults

Table O.44: Olanzapine versus risperidone in adults

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With risperidone	With olanzapine		Risk with risperidone	Risk difference with olanzapine (95% CI)
Targeted behaviour that challenges (frequency) – post-treatment (Better indicated by lower values)											
62 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	31	31	-		The mean targeted behaviour that challenges (frequency) – post-treatment in the intervention groups was 0.2 standard deviations higher (0.3 lower to 0.7 higher)
Adverse events (elevated prolactin) – post-treatment											
62 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk	30/31 (96.8%)	22/31 (71%)	RR 0.73 (0.58 to	968 per 1000	261 fewer per 1000 (from 68 fewer to 406 fewer)

						of bias, imprecision			0.93)		
Adverse events (weight gain, non-occurrence) – post-treatment											
62 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⊕⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	28/31 (90.3%)	24/31 (77.4%)	RR 0.86 (0.69 to 1.07)	903 per 1000	126 fewer per 1000 (from 280 fewer to 63 more)
Adverse events (sedation, non-occurrence) – post-treatment											
62 (1 study)	serious[2]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⊕⊖⊖⊖ VERY LOW[2] due to risk of bias, imprecision	26/31 (83.9%)	24/31 (77.4%)	RR 0.92 (0.72 to 1.18)	839 per 1000	67 fewer per 1000 (from 235 fewer to 151 more)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Optimal information size not met; small, single study

A.9.17 Lithium versus placebo in adults

Table O.45: Lithium versus placebo in adults

Quality assessment							**Summary of findings**				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With placebo	With lithium		Risk with placebo	Risk difference with lithium (95% CI)
Targeted behaviour that challenges (frequency, non-improvement)											
42 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⊕⊖⊖⊖ VERY LOW[1,2]	14/20 (70%)	6/22 (27.3	RR 0.39 (0.19 to	700 per 1000	427 fewer per 1000

Quality assessment							Summary of findings				
						due to risk of bias, imprecision		%)	0.82)		(from 126 fewer to 567 fewer)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Optimal information size not met; small, single study

A.9.18 Withdrawal of zuclopenthixol versus continuation of zuclopenthixol in adults

Table O.46: Withdrawal of zuclopenthixol versus continuation of zuclopenthixol in adults

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With continuation of zuclopenthixol	With withdrawal of zuclopenthixol		Risk with continuation of zuclopenthixol	Risk difference with withdrawal of zuclopenthixol (95% CI)
Targeted behaviour that challenges (relapse) – post-treatment											
39 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	12/19 (63.2%)	19/20 (95%)	RR 1.5 (1.05 to 2.15)	632 per 1000	316 more per 1000 (from 32 more to 726 more)
Targeted behaviour that challenges (severity) – post-treatment (measured with: End-point score; Better indicated by lower values)											
39 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	19	20	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.56 standard deviations higher

											(0.08 lower to 1.2 higher)
Targeted behaviour that challenges (severity) – post-treatment (measured with: Change score; Better indicated by lower values)											
85 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	45	40	-		The mean targeted behaviour that challenges (severity) – post-treatment in the intervention groups was 0.68 standard deviations higher (0.24 to 1.11 higher)
Targeted behaviour that challenges (problems in management) – post-treatment											
43 (1 study)	serious[3]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[2,3] due to risk of bias, imprecision	5/24 (20.8%)	7/19 (36.8%)	RR 1.77 (0.67 to 4.7)	208 per 1000	160 more per 1000 (from 69 fewer to 771 more)
Adaptive functioning (social) – post-treatment (Better indicated by higher values)											
85 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	45	40	-		The mean adaptive functioning (social) – post-treatment in the intervention groups was 0.47 standard deviations lower (0.9 to 0.04 lower)
Adverse events (weight gain; kg) – post- treatment (Better indicated by lower values)											
39 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	19	20	-		The mean adverse events (weight gain; kg) – post- treatment in the intervention groups was 0.55 standard

											deviations lower (1.19 lower to 0.09 higher)
Adverse events (drowsiness, non-occurrence) – post-treatment											
42 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	19/20 (95%)	21/22 (95.5%)	RR 1 (0.88 to 1.15)	950 per 1000	0 fewer per 1000 (from 114 fewer to 142 more)
Adverse events (discontinuation due to adverse events, non-occurrence) – post-treatment											
204 (3 studies)	serious[4]	serious[5]	no serious indirectness	serious[6]	undetected	⨁⊖⊖⊖ VERY LOW[4,5,6] due to risk of bias, inconsistency, imprecision	98/103 (95.1%)	80/101 (79.2%)	RR 0.86 (0.71 to 1.04)	951 per 1000	133 fewer per 1000 (from 276 fewer to 38 more)
Adverse events (discontinuation due to other reasons, non-occurrence) – post-treatment											
91 (2 studies)	serious[4]	very serious[7]	no serious indirectness	serious[6]	undetected	⨁⊖⊖⊖ VERY LOW[4,6,7] due to risk of bias, inconsistency, imprecision	38/46 (82.6%)	29/45 (64.4%)	RR 0.73 (0.33 to 1.64)	826 per 1000	223 fewer per 1000 (from 553 fewer to 529 more)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Optimal information size not met; small, single study
[3] Crucial limitation for one or more criteria sufficient to substantially lower ones confidence in the estimate of effect
[4] Most information is from studies at moderate risk of bias
[5] $I^2 > 40\%$
[6] Optimal information size not met
[7] $I^2 > 75\%$

A.9.19 Melatonin versus placebo in children and young people

Table O.47: Melatonin versus placebo in children and young people

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%) With placebo	Study event rates (%) With melatonin	Relative effect (95% CI)	Anticipated absolute effects Risk with placebo	Anticipated absolute effects Risk difference with melatonin (95% CI)
Targeted behaviour that challenges (global problem sleep behaviour) – post-treatment (measured with: Children's Sleep Habits Questionnaire; Better indicated by lower values)											
66 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	32	34	-		The mean targeted behaviour that challenges (global problem sleep behaviour) – post-treatment in the intervention groups was 1.81 standard deviations lower (2.39 to 1.23 lower)
Targeted behaviour that challenges (global problem sleep behaviour) – post-treatment (measured with: Composite Sleep Disturbance Index; Better indicated by lower values)											
125 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⨁⊖⊖ LOW[3] due to imprecision	65	60	-		The mean targeted behaviour that challenges (global problem sleep behaviour) – post-treatment in the intervention groups was 0.26 standard deviations lower (0.62 lower to 0.09 higher)
Targeted behaviour that challenges (non-improvement of global problem sleep behaviour) – post-treatment											
66 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of	32/32 (100 %)	21/34 (61.8 %)	RR 0.62 (0.48	1000 per 1000	380 fewer per 1000 (from 190 fewer to 520 fewer)

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%) With placebo	Study event rates (%) With melatonin	Relative effect (95% CI)	Anticipated absolute effects Risk with placebo	Anticipated absolute effects Risk difference with melatonin (95% CI)
Targeted behaviour that challenges (global problem sleep behaviour) – post-treatment (measured with: Children's Sleep Habits Questionnaire; Better indicated by lower values)											
66 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	32	34	-		The mean targeted behaviour that challenges (global problem sleep behaviour) – post-treatment in the intervention groups was 1.81 standard deviations lower (2.39 to 1.23 lower)
						bias, indirectness, imprecision			to 0.81)		
Targeted behaviour that challenges (sleep efficiency) – post-treatment (measured with: Actigraph; Better indicated by higher values)											
124 (2 studies)	no serious risk of bias	very serious[4]	no serious indirectness	serious[5]	undetected	⨁⊖⊖⊖ VERY LOW[4,5] due to inconsistency, imprecision	60	64	-		The mean targeted behaviour that challenges (sleep efficiency) – post-treatment in the intervention groups was 1.46 standard deviations higher (0.51 lower to 3.42 higher)
Targeted behaviour that challenges (total sleep time) – post-treatment (measured with: Actigraph; Better indicated by higher values)											
125 (2 studies)	no serious risk of bias	very serious[4]	no serious indirectness	serious[5]	undetected	⨁⊖⊖⊖ VERY LOW[4,5] due to inconsistency, imprecision	61	64	-		The mean targeted behaviour that challenges (total sleep time) – post-treatment in the intervention groups was 1.01 standard deviations higher (0.26 lower to 2.28 higher)

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%) With placebo	Study event rates (%) With melatonin	Relative effect (95% CI)	Anticipated absolute effects Risk with placebo	Anticipated absolute effects Risk difference with melatonin (95% CI)
Targeted behaviour that challenges (global problem sleep behaviour) – post-treatment (measured with: Children's Sleep Habits Questionnaire; Better indicated by lower values)											
66 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⨀⨀⨀ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	32	34	-		The mean targeted behaviour that challenges (global problem sleep behaviour) – post-treatment in the intervention groups was 1.81 standard deviations lower (2.39 to 1.23 lower)
Targeted behaviour that challenges (wake after sleep onset) – post-treatment (measured with: Actigraph; Better indicated by lower values)											
115 (2 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[5]	undetected	⨁⨁⨁⨀ MODERATE[5] due to imprecision	57	58	-		The mean targeted behaviour that challenges (wake after sleep onset) – post-treatment in the intervention groups was 0.76 standard deviations lower (1.14 to 0.38 lower)
Targeted behaviour that challenges (sleep onset latency) – post-treatment (measured with: Actigraph; Better indicated by lower values)											
66 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⨀⨀⨀ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	32	34	-		The mean targeted behaviour that challenges (sleep onset latency) – post-treatment in the intervention groups was 1.23 standard deviations lower (1.75 to 0.7 lower)
Targeted behaviour that challenges (total sleep time) – post-treatment (measured with: Sleep diary; Better indicated by higher values)											
169 (3	no serious	serious[6]	no serious indirectnes	serious[5]	undetected	⨁⨁⨀⨀ LOW[5,6]	85	84	-		The mean targeted behaviour that challenges (total sleep

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%) With placebo	Study event rates (%) With melatonin	Relative effect (95% CI)	Anticipated absolute effects Risk with placebo	Anticipated absolute effects Risk difference with melatonin (95% CI)
Targeted behaviour that challenges (global problem sleep behaviour) – post-treatment (measured with: Children's Sleep Habits Questionnaire; Better indicated by lower values)											
66 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⨀⨀⨀ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	32	34	-		The mean targeted behaviour that challenges (global problem sleep behaviour) – post-treatment in the intervention groups was 1.81 standard deviations lower (2.39 to 1.23 lower)
studies)	risk of bias		s			due to inconsistency, imprecision					time) – post-treatment in the intervention groups was 0.34 standard deviations higher (0.37 lower to 1.05 higher)
Targeted behaviour that challenges (number of wakes per night) – post-treatment (measured with: Sleep diary; Better indicated by lower values)											
164 (3 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[5]	undetected	⨁⨁⨁⨀ MODERATE[5] due to imprecision	81	83	-		The mean targeted behaviour that challenges (number of wakes per night) – post-treatment in the intervention groups was 0.06 standard deviations lower (0.49 lower to 0.37 higher)
Targeted behaviour that challenges (wake after sleep onset) – post-treatment (measured with: Sleep diary; Better indicated by lower values)											
172 (3 studies)	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[5]	undetected	⨁⨁⨁⨀ MODERATE[5] due to imprecision	85	87	-		The mean targeted behaviour that challenges (wake after sleep onset) – post-treatment in the intervention groups was

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%) With placebo	Study event rates (%) With melatonin	Relative effect (95% CI)	Anticipated absolute effects Risk with placebo	Anticipated absolute effects Risk difference with melatonin (95% CI)
Targeted behaviour that challenges (global problem sleep behaviour) – post-treatment (measured with: Children's Sleep Habits Questionnaire; Better indicated by lower values)											
66 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⨀⨀⨀ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	32	34	-		The mean targeted behaviour that challenges (global problem sleep behaviour) – post-treatment in the intervention groups was 1.81 standard deviations lower (2.39 to 1.23 lower)
											0.64 standard deviations lower (1.03 to 0.25 lower)
Targeted behaviour that challenges (duration of wakes) – post-treatment (measured with: Sleep diary; Better indicated by lower values)											
163 (3 studies)	no serious risk of bias	serious[6]	no serious indirectness	serious[5]	undetected	⨁⨁⨀⨀ LOW[5,6] due to inconsistency, imprecision	81	82	-		The mean targeted behaviour that challenges (duration of wakes) – post-treatment in the intervention groups was 0.23 standard deviations higher (0.36 lower to 0.82 higher)
Adverse events (solomnence/sedation, non-occurrence) – post-treatment											
146 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⨁⨀⨀ LOW[3] due to imprecision	66/76 (86.8 %)	61/70 (87.1 %)	RR 1 (0.89 to 1.14)	868 per 1000	0 fewer per 1000 (from 96 fewer to 122 more)
Adverse events (discontinuation due to adverse events, non-occurrence) – post-treatment											
146 (1 study)	no serious	no serious inconsistenc	no serious indirectnes	very serious[3]	undetected	⨁⨁⨀⨀ LOW[3]	74/76 (97.4	69/70 (98.6	RR 1.01	974 per	10 more per 1000 (from 29 fewer to 58 more)

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%) With placebo	Study event rates (%) With melatonin	Relative effect (95% CI)	Anticipated absolute effects Risk with placebo	Anticipated absolute effects Risk difference with melatonin (95% CI)
Targeted behaviour that challenges (global problem sleep behaviour) – post-treatment (measured with: Children's Sleep Habits Questionnaire; Better indicated by lower values)											
66 (1 study)	serious[1] risk of bias	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⨀⨀⨀ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	32 %)	34 %)	- (0.97 to 1.06)	1000	The mean targeted behaviour that challenges (global problem sleep behaviour) – post-treatment in the intervention groups was 1.81 standard deviations lower (2.39 to 1.23 lower)
Adverse events (discontinuation due to other reasons, non-occurrence) – post-treatment											
284 (3 studies)	no serious risk of bias	serious[6]	no serious indirectness	serious[5]	undetected	⨁⨁⨀⨀ LOW[5,6] due to inconsistency, imprecision	127/144 (88.2 %)	130/140 (92.9 %)	RR 1.06 (0.94 to 1.2)	882 per 1000	53 more per 1000 (from 53 fewer to 176 more)
Adverse events (seizure, non-occurrence) – post-treatment											
146 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⨁⨀⨀ LOW[3] due to imprecision	75/76 (98.7 %)	70/70 (100 %)	RR 1.01 (0.98 to 1.05)	987 per 1000	10 more per 1000 (from 20 fewer to 49 more)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Applicability- different populations

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%) With placebo	Study event rates (%) With melatonin	Relative effect (95% CI)	Anticipated absolute effects Risk with placebo	Anticipated absolute effects Risk difference with melatonin (95% CI)
Targeted behaviour that challenges (global problem sleep behaviour) – post-treatment (measured with: Children's Sleep Habits Questionnaire; Better indicated by lower values)											
66 (1 study)	serious [1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	32	34	-		The mean targeted behaviour that challenges (global problem sleep behaviour) – post-treatment in the intervention groups was 1.81 standard deviations lower (2.39 to 1.23 lower)

[3] Optimal information size not met; small, single study
[4] I^2 > 75%
[5] Optimal information size not met
[6] I^2 > 40%

A.9.20 Melatonin versus cognitive behavioural therapy in children and young people

Table O.48: Melatonin versus cognitive behavioural therapy in children and young people

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With CBT	With melatonin		Risk with CBT	Risk difference with melatonin (95% CI)
Targeted behaviour that challenges (global problem sleep behaviour) – post-treatment (measured with: Children's Sleep Habits Questionnaire; Better indicated by lower values)											
67 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	33	34	-		The mean targeted behaviour that challenges (global problem sleep behaviour) – post-treatment in the intervention groups was 0.94 standard deviations lower (1.45 to 0.44 lower)
Targeted behaviour that challenges (non-improvement of global sleep problem behaviour) – post-treatment											
67 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	30/33 (90.9%)	21/34 (61.8%)	RR 0.68 (0.51 to 0.9)	909 per 1000	291 fewer per 1000 (from 91 fewer to 445 fewer)
Targeted behaviour that challenges (sleep onset latency) – post-treatment (measured with: Actigraph; Better indicated by lower values)											
67 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	33	34	-		The mean targeted behaviour that challenges (sleep onset latency) – post-treatment in the intervention groups was 0.54 standard deviations lower (1.03 to 0.05 lower)

Targeted behaviour that challenges (wake after sleep onset) – post-treatment (measured with: Actigraph; Better indicated by lower values)											
67 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	33	34	-		The mean targeted behaviour that challenges (wake after sleep onset) – post-treatment in the intervention groups was 0.73 standard deviations lower (1.22 to 0.23 lower)
Targeted behaviour that challenges (total sleep time) – post-treatment (measured with: Actigraph; Better indicated by higher values)											
67 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	33	34	-		The mean targeted behaviour that challenges (total sleep time) – post-treatment in the intervention groups was 0.76 standard deviations higher (0.26 to 1.26 higher)
Targeted behaviour that challenges (sleep efficiency) – post-treatment (measured with: Actigraph; Better indicated by higher values)											
67 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	33	34	-		The mean targeted behaviour that challenges (sleep efficiency) – post-treatment in the intervention groups was 0.89 standard deviations higher (0.39 to 1.4 higher)
Adverse events (discontinuation due to other reasons, non-occurrence) – post-treatment											
80 (1 study)	serious[1]	no serious inconsistency	serious[2]	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	36/40 (90 %)	36/40 (90%)	RR 1 (0.86 to 1.16)	900 per 1000	0 fewer per 1000 (from 126 fewer to 144 more)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Applicability- different populations
[3] Optimal information size not met; small, single study

A.10 Interventions aimed at improving the health and well-being of carers of people with learning disabilities

A.10.1 Cognitive behavioural interventions for family carers versus any control

Table O.49: Cognitive behavioural interventions for family carers versus any control

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With Any control	With cognitive behavioural intervention		Risk with any control	Risk difference with cognitive behavioural intervention (95% CI)
Carer health and well-being (depression) – post-treatment (Better indicated by lower values)											
428 (5 studies)	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	undetected	⨁⨁⊖⊖ LOW[1,2] due to risk of bias, imprecision	177	251	-		The mean carer health and well-being (depression) – post-treatment in the intervention groups was 0.35 standard deviations lower (0.54 to 0.15 lower)
Carer health and well-being (depression) – follow-up (Better indicated by lower values)											
130 (2 studies) 46 to 104 weeks	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	undetected	⨁⨁⊖⊖ LOW[1,2] due to risk of bias, imprecision	66	64	-		The mean carer health and well-being (depression) – follow-up in the intervention groups was 0.41 standard deviations lower (0.79 to 0.04 lower)

Quality assessment							Summary of findings				
Carer health and well-being (clinically depressed) – post-treatment											
111 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[1,3] due to risk of bias, imprecision	13/58 (22.4 %)	3/53 (5.7%)	RR 0.25 (0.08 to 0.84)	224 per 1000	168 fewer per 1000 (from 36 fewer to 206 fewer)
Carer health and well-being (anxiety, trait) – post-treatment (Better indicated by lower values)											
68 (2 studies)	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	undetected	⨁⨁⊖⊖ LOW[1,2] due to risk of bias, imprecision	31	37	-		The mean carer health and well-being (anxiety, trait) – post-treatment in the intervention groups was 0.5 standard deviations lower (1.03 lower to 0.03 higher)
Carer health and well-being (anxiety, state) – post-treatment (Better indicated by lower values)											
36 (1 study)	serious[4]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[3,4] due to risk of bias, imprecision	18	18	-		The mean carer health and well-being (anxiety, state) – post-treatment in the intervention groups was 0.46 standard deviations lower (1.12 lower to 0.2 higher)
Carer health and well-being (mental ill health) – post-treatment (Better indicated by lower values)											
58 (1 study)	serious[4]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[3,4] due to risk of bias, imprecision	29	29	-		The mean carer health and well-being (mental ill health) – post-treatment in the intervention groups was 2.19 standard deviations lower (2.85 to 1.53 lower)

Quality assessment							Summary of findings				
Carer health and well-being (quality of life) – post-treatment (Better indicated by lower values)											
58 (1 study)	serious[4]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[3,4] due to risk of bias, imprecision	29	29	-		The mean carer health and well-being (quality of life) – post-treatment in the intervention groups was 0.87 standard deviations higher (0.33 to 1.41 higher)
Carer health and well-being (stress) – post-treatment (Better indicated by lower values)											
384 (3 studies)	serious[1]	serious[5]	no serious indirectness	serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2,5] due to risk of bias, inconsistency, imprecision	159	225	-		The mean carer health and well-being (stress) – post-treatment in the intervention groups was 0.45 standard deviations lower (0.78 to 0.12 lower)
Carer health and well-being (stress) – follow-up (Better indicated by lower values)											
76 (1 study) 104 weeks	serious[4]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[3,4] due to risk of bias, imprecision	27	49	-		The mean carer health and well-being (stress) – follow-up in the intervention groups was 0.43 standard deviations lower (0.9 lower to 0.05 higher)
Carer health and well-being (clinically stressed) – post-treatment											
111 (1 study)	serious[4]	no serious inconsistency	no serious indirectness	very serious[3]	undetected	⨁⊖⊖⊖ VERY LOW[3,4] due to risk of bias, imprecision	17/58 (29.3 %)	2/53 (3.8%)	RR 0.13 (0.03 to 0.53)	293 per 1000	255 fewer per 1000 (from 138 fewer to 284 fewer)

1 Most information is from studies at moderate risk of bias
[2] Optimal information size not met

Quality assessment	Summary of findings
[3] Optimal information size not met; small, single study [4] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect [5] I^2 > 40%	

A.10.2 Psychoeducational interventions for family carers versus any control

Table O.50: Psychoeducational interventions for family carers versus any control

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With any control	With psychoeducation		Risk with any control	Risk difference with psychoeducation (95% CI)
Carer health and well-being (depression) – follow-up (Better indicated by lower values)											
75 (1 study) 4 weeks	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	35	40	-		The mean carer health and well-being (depression) – follow-up in the intervention groups was 0.84 standard deviations lower (1.31 to 0.36 lower)
Carer health and well-being (burnout) – follow-up (Better indicated by lower values)											
90 (1 study) 8 weeks	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias,	45	45	-		The mean carer health and well-being (burnout) – follow-up in the intervention groups was 0.35 standard deviations

Quality assessment							Summary of findings				
						imprecision					lower (0.77 lower to 0.06 higher)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Optimal information size not met; small, single study

A.10.3 Support interventions for family carers versus any control

Table O.51: Parent advisor scheme versus treatment as usual

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With any control	With support interventions		Risk with any control	Risk difference with support interventions (95% CI)
Carer health and well-being (stress) – post-treatment (Better indicated by lower values)											
28 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	12	16	-		The mean carer health and well-being (stress) – post-treatment in the intervention groups was 1.21 standard deviations lower (2.04 to 0.39 lower)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Optimal information size not met; small, single study

A.10.4 Mindfulness interventions for paid carers versus any control

Table O.52: Mindfulness interventions for paid carers versus any control

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With any control	With mindfulness interventions		Risk with any control	Risk difference with mindfulness interventions (95% CI)
Carer health and well-being (mental well-being) – post-treatment (Better indicated by higher values)											
120 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	54	66	-		The mean carer health and well-being (mental well-being) – post-treatment in the intervention groups was 0.17 standard deviations higher (0.19 lower to 0.53 higher)
Carer health and well-being (mental well-being) – follow-up (Better indicated by higher values)											
120 (1 study) 6 weeks	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	54	66	-		The mean carer health and well-being (mental well-being) – follow-up in the intervention groups was 0.28 standard deviations higher (0.08 lower to 0.64 higher)
Carer health and well-being (mental ill health) – post-treatment (Better indicated by lower values)											
154 (2 studies)	serious[3]	serious[4]	no serious indirectness	serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[2,3,4] due to risk of bias,	70	84	-		The mean carer health and well-being (mental ill health) – post-treatment in the intervention groups was

Quality assessment							Summary of findings				
						inconsistency, imprecision					0.54 standard deviations lower (1.06 to 0.02 lower)
Carer health and well-being (mental ill health) – follow-up (Better indicated by lower values)											
154 (2 studies) 6-13 weeks	serious[3]	serious[4]	no serious indirectness	serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[2,3,4] due to risk of bias, inconsistency, imprecision	70	84	-		The mean carer health and well-being (mental ill health) – follow-up in the intervention groups was 0.24 standard deviations lower (0.72 lower to 0.24 higher)
Carer health and well-being (stress) – post-treatment (Better indicated by lower values)											
120 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	54	66	-		The mean carer health and well-being (stress) – post-treatment in the intervention groups was 0.17 standard deviations higher (0.19 lower to 0.53 higher)
Carer health and well-being (stress) – follow-up (Better indicated by lower values)											
120 (1 study) 6 weeks	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	54	66	-		The mean carer health and well-being (stress) – follow-up in the intervention groups was 0.05 standard deviations lower (0.41 lower to 0.31 higher)
Carer health and well-being (burnout) – post-treatment (Better indicated by lower values)											
34 (1 study)	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias,	16	18	-		The mean carer health and well-being (burnout) – post-treatment in the intervention groups was

Quality assessment							Summary of findings				
						imprecision					0.18 standard deviations lower (0.86 lower to 0.49 higher)
Carer health and well-being (burnout) – follow-up (Better indicated by lower values)											
34 (1 study) 13 weeks	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	undetected	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	16	18	-		The mean carer health and well-being (burnout) – follow-up in the intervention groups was 0.08 standard deviations lower (0.76 lower to 0.59 higher)

[1] Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect
[2] Optimal information size not met; small, single study
3 Most information is from studies at moderate risk of bias
[4] I^2 > 40%