

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

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

; RR: Risk ratio; MD: Mean difference

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