GRADE tables for review question: What are the relative benefits and harms of further-line psychological, psychosocial, pharmacological and physical interventions (alone or in combination), for adults with depression showing an inadequate response to at least one previous intervention for the current episode?

Table 70: Clinical evidence profile for comparison 1. Augmenting with cognitive and cognitive behavioural therapies versus continuing with antidepressant (+/ waitlist or attention-placebo)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Augmenting with cognitive and cognitive behavioural therapies	Continuing with antidepressant (+/ waitlist or attention-placebo)	Relative (95% CI)	Absolute		
Depression symptomatology endpoint (follow-up 8-26 weeks; measured with: Beck Depression Inventory (BDI/BDI-II) or Hamilton Rating Scale for Depression (HAMD); Better indicated by lower values)												
13 (Chan 2012, Chiesa 2015, Dozois 2009, Dunn 1979, Embling 2002, Kocsis 2009/ Klein 2011, Lynch 2007_study 2, Nakagawa 2017, Nakao 2018, Paykel 1999/ Scott 2000, Strauss 2012, Watkins 2011a, Wiles 2013/2016)	randomised trials	serious[1]	serious[2]	no serious indirectness	serious[3]	none	666	558	-	SMD 0.74 lower (1.03 to 0.45 lower)	VERY LOW	CRITICAL
Depression symptomatology change score (follow-up 8-26 weeks; measured with: Beck Depression Inventory (BDI/BDI-II) or Hamilton Rating Scale for Depression (HAMD) change from baseline to endpoint; Better indicated by lower values)												
10 (Chan 2012, Chiesa 2015, Dozois 2009, Dunn 1979, Embling 2002, Nakagawa 2017, Nakao 2018, Paykel 1999/ Scott 2000, Strauss 2012, Watkins 2011a)	randomised trials	serious[1]	very serious[4]	no serious indirectness	no serious imprecision	none	265	259	-	SMD 1.36 lower (1.87 to 0.86 lower)	VERY LOW	CRITICAL
Depression symptomatology at 2-3 month follow-up (follow-up 8-16 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												

2 (Chiesa 2015, Nakagawa 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	63	60	-	SMD 0.51 lower (0.87 to 0.15 lower)	MODERATE	CRITICAL
Depression symptomatology at 4-6 month follow-up (follow-up mean 4-6 months; measured with: Hamilton Rating Scale for Depression (HAM-D)/Beck Depression Inventory (BDI/BDI-II); Better indicated by lower values)												
5 (Chiesa 2015, Dunn 1979, Nakagawa 2017, Paykel 1999/ Scott 2000, Wiles 2013/2016)	randomised trials	no serious risk of bias	serious[2]	no serious indirectness	serious[3]	none	350	346	-	SMD 0.51 lower (0.77 to 0.24 lower)	LOW	CRITICAL
Depression symptomatology at 11-12 month follow-up (follow-up 11-12 months; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
2 (Nakagawa 2017, Paykel 1999/ Scott 2000)	randomised trials	no serious risk of bias	very serious[4]	no serious indirectness	serious[3]	none	120	118	-	SMD 0.3 lower (0.93 lower to 0.33 higher)	VERY LOW	CRITICAL
Depression symptomatology at 40-month follow-up (follow-up mean 40 months; measured with: Beck Depression Inventory (BDI-II); Better indicated by lower values)												
1 (Wiles 2013/2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	136	112	-	SMD 0.31 lower (0.56 to 0.06 lower)	LOW	CRITICAL
Remission (ITT) (follow-up 8-26 weeks; assessed with: Number of people scoring =<7/10 on Hamilton Rating Scale for Depression (HAM-D) or <10 on Beck Depression Inventory (BDI-II))												
8 (Eisendrath 2016, Kocsis 2009/ Klein 2011, Lynch 2007_study 2, Nakagawa 2017, Nakao 2018, Paykel 1999/ Scott 2000, Watkins 2011a, Wiles 2013/2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	215/703 (30.6%)	101/590 (17.1%)	RR 1.76 (1.32 to 2.36)	130 more per 1000 (from 55 more to 233 more)	MODERATE	CRITICAL
Remission (ITT) at 3-month follow-up (follow-up mean 3 months; assessed with: Number of people scoring =<7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Nakagawa 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	20/40 (50%)	12/40 (30%)	RR 1.67 (0.95 to 2.93)	201 more per 1000 (from 15 fewer to 579 more)	MODERATE	CRITICAL
Remission (ITT) at 6-month follow-up (follow-up mean 6 months; assessed with: Number of people scoring <10 on Beck Depression Inventory (BDI-II)/≤7 on Hamilton Rating Scale for Depression (HAM-D))												
2 (Nakagawa 2017, Wiles 2013/2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	106/274 (38.7%)	52/275 (18.9%)	RR 1.99 (1.52 to 2.62)	187 more per 1000 (from 98 more to 306 more)	MODERATE	CRITICAL

Remission (ITT) at 12-month follow-up (follow-up mean 12 months; assessed with: Number of people scoring =<7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Nakagawa 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	29/40 (72.5%)	17/40 (42.5%)	RR 1.71 (1.13 to 2.56)	302 more per 1000 (from 55 more to 663 more)	MODERATE	CRITICAL
Remission (ITT) at 40-month follow-up (follow-up mean 40 months; assessed with: Number of people scoring <10 on Beck Depression Inventory (BDI-II))												
1 (Wiles 2013/2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	38/234 (16.2%)	20/235 (8.5%)	RR 1.91 (1.15 to 3.18)	77 more per 1000 (from 13 more to 186 more)	LOW	CRITICAL
Response (ITT) (follow-up 8-26 weeks; assessed with: Response: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D)/Beck Depression Inventory (BDI-II))												
6 (Eisendrath 2016, Nakagawa 2017, Nakao 2018, Watkins 2011a, Wiles 2008, Wiles 2013/2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	189/416 (45.4%)	81/413 (19.6%)	RR 2.27 (1.83 to 2.83)	249 more per 1000 (from 163 more to 359 more)	MODERATE	CRITICAL
Response (ITT) at 3-month follow-up (follow-up mean 3 months; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (Nakagawa 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	28/40 (70%)	17/40 (42.5%)	RR 1.65 (1.09 to 2.49)	276 more per 1000 (from 38 more to 633 more)	MODERATE	CRITICAL
Response (ITT) at 6-month follow-up (follow-up mean 6 months; assessed with: Number of people showing at least 50% improvement on Beck Depression Inventory (BDI-II)/Hamilton Rating Scale for Depression (HAM-D))												
2 (Nakagawa 2017, Wiles 2013/2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	143/274 (52.2%)	86/275 (31.3%)	RR 1.6 (1.27 to 2.01)	188 more per 1000 (from 84 more to 316 more)	MODERATE	CRITICAL
Response (ITT) at 12-month follow-up (follow-up mean 12 months; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (Nakagawa 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	33/40 (82.5%)	20/40 (50%)	RR 1.65 (1.17 to 2.32)	325 more per 1000 (from 85 more to 660 more)	MODERATE	CRITICAL
Response (ITT) at 40-month follow-up (follow-up mean 40 months; assessed with: Number of people showing at least 50% improvement on Beck Depression Inventory (BDI-II))												

1 (Wiles 2013/2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	59/234 (25.2%)	30/235 (12.8%)	RR 1.98 (1.32 to 2.95)	125 more per 1000 (from 41 more to 249 more)	MODERATE	CRITICAL
Discontinuation due to any reason (follow-up 8-26 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
13 (Chan 2012, Chiesa 2015, Dozois 2009, Eisendrath 2016, Kocsis 2009/ Klein 2011, Lynch 2007_study 2, Nakagawa 2017, Nakao 2018, Paykel 1999/ Scott 2000, Strauss 2012, Watkins 2011a, Wiles 2008, Wiles 2013/2016)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	111/807 (13.8%)	103/687 (15%)	RR 0.95 (0.74 to 1.21)	7 fewer per 1000 (from 39 fewer to 31 more)	MODERATE	CRITICAL
Discontinuation due to side effects (follow-up mean 12 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Kocsis 2009/ Klein 2011)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[5]	none	2/200 (1%)	2/96 (2.1%)	RR 0.48 (0.07 to 3.36)	11 fewer per 1000 (from 19 fewer to 49 more)	LOW	CRITICAL
Quality of life endpoint (follow-up mean 12 weeks; measured with: European Quality of Life Questionnaire-5 Dimensions (EQ-5D); Better indicated by higher values)												
1 (Nakao 2018)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[5]	none	20	20	-	SMD 0 higher (0.62 lower to 0.62 higher)	LOW	IMPORTANT
Quality of life physical component score (PCS) endpoint (follow-up 12-26 weeks; measured with: 12-item/36-item Short-Form Survey (SF-12/SF-36): Physical component score; Better indicated by higher values)												
3 (Nakagawa 2017, Nakao 2018, Wiles 2013/2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	261	269	-	SMD 0.04 higher (0.17 lower to 0.26 higher)	MODERATE	IMPORTANT
Quality of life mental component score (MCS) endpoint (follow-up 12-26 weeks; measured with: 12-item/36-item Short-Form Survey (SF-12/SF-36): Mental component score; Better indicated by higher values)												
3 (Nakagawa 2017, Nakao 2018, Wiles 2013/2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	261	269	-	SMD 0.26 higher (0.03 lower to 0.55 higher)	LOW	IMPORTANT
Quality of life physical component score (PCS) at 3-month follow-up (follow-up mean 3 months; measured with: 36-item Short-Form Survey (SF-36): Physical component score; Better indicated by higher values)												
1 (Nakagawa 2017)	randomised trials	no serious	no serious inconsistency	no serious indirectness	serious[3]	none	40	40	-	SMD 0.17 lower (0.61	MODERATE	IMPORTANT

		risk of bias								lower to 0.27 higher)		
Quality of life mental component score (MCS) at 3-month follow-up (follow-up mean 3 months; measured with: 36-item Short-Form Survey (SF-36): Mental component score; Better indicated by higher values)												
1 (Nakagawa 2015)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[6]	none	40	40	-	SMD 0.15 lower (0.58 lower to 0.29 higher)	MODERATE	IMPORTANT
Quality of life physical component score (PCS) at 6-month follow-up (follow-up mean 6 months; measured with: 12-item/36-item Short-Form Survey (SF-12/SF-36): Physical component score; Better indicated by higher values)												
2 (Nakagawa 2015, Wiles 2013/2016)	randomised trials	serious[1]	serious[2]	no serious indirectness	serious[3]	none	234	235	-	SMD 0.07 higher (0.37 lower to 0.52 higher)	VERY LOW	IMPORTANT
Quality of life mental component score (MCS) at 6-month follow-up (follow-up mean 6 months; measured with: 12-item/36-item Short-Form Survey (SF-12/SF-36): Mental component score; Better indicated by higher values)												
2 (Nakagawa 2015, Wiles 2013/2016)	randomised trials	serious[1]	very serious[4]	no serious indirectness	very serious[5]	none	234	235	-	SMD 0.01 higher (0.56 lower to 0.58 higher)	VERY LOW	IMPORTANT
Quality of life physical component score (PCS) at 12-month follow-up (follow-up mean 12 months; measured with: 36-item Short-Form Survey (SF-36): Physical component score; Better indicated by higher values)												
1 (Nakagawa 2015)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	40	40	-	SMD 0.05 higher (0.39 lower to 0.49 higher)	HIGH	IMPORTANT
Quality of life mental component score (MCS) at 12-month follow-up (follow-up mean 12 months; measured with: 36-item Short-Form Survey (SF-36): Mental component score; Better indicated by higher values)												
1 (Nakagawa 2015)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[6]	none	40	40	-	SMD 0.2 lower (0.64 lower to 0.24 higher)	MODERATE	IMPORTANT
Quality of life physical component score (PCS) at 40-month follow-up (follow-up mean 40 months; measured with: 12-item Short-Form Survey (SF-12): Physical component score; Better indicated by higher values)												
1 (Wiles 2013/2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	132	110	-	SMD 0.22 higher (0.03 lower to 0.47 higher)	MODERATE	IMPORTANT
Quality of life mental component score (MCS) at 40-month follow-up (follow-up mean 40 months; measured with: 12-item Short-Form Survey (SF-12): Mental component score; Better indicated by higher values)												
1 (Wiles 2013/2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	132	110	-	SMD 0.34 higher (0.09	LOW	IMPORTANT

										to 0.6 higher)		
Functional impairment endpoint (follow-up 12-20 weeks; measured with: Longitudinal Interval Follow-up Evaluation Range of Impaired Functioning Tool (LIFE-RIFT)/Social Adjustment Scale (SAS); Better indicated by lower values)												
2 (Kocsis 2009/ Klein 2011, Paykel 1999/ Scott 2000)	randomised trials	no serious risk of bias	serious[2]	no serious indirectness	serious[3]	none	252	153	-	SMD 0.36 lower (0.67 to 0.05 lower)	LOW	IMPORTANT
Functional impairment at 11-month follow-up (follow-up mean 11 months; measured with: Social Adjustment Scale (SAS); Better indicated by lower values)												
1 (Paykel 1999/ Scott 2000)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	80	78	-	SMD 0.3 lower (0.61 lower to 0.01 higher)	MODERATE	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains
[2] Substantial heterogeneity
[3] 95% CI crosses threshold for both clinically important benefit and no effect
[4] Considerable heterogeneity
[5] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[6] 95% CI crosses threshold for both clinically important harm and no effect

Table 71: Clinical evidence profile for comparison 2. Augmenting with cognitive and cognitive behavioural therapies versus augmenting with counselling

Quality assessment							**No of patients**		**Effect**		**Quality**	**Importance**
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with cognitive and cognitive behavioural therapies**	**Augmenting with counselling**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 12 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Kocsis 2009/ Klein 2011)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	174	168	-	SMD 0.18 lower (0.39 lower to 0.04 higher)	HIGH	CRITICAL
Remission (ITT) (follow-up mean 12 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D) AND responding (≥50% improvement on HAM-D))												
1 (Kocsis 2009/ Klein 2011)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	67/200 (33.5%)	52/195 (26.7%)	RR 1.26 (0.93 to 1.7)	69 more per 1000 (from 19 fewer to 187 more)	MODERATE	CRITICAL

Discontinuation due to any reason (follow-up mean 12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Kocsis 2009/ Klein 2011)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	25/200 (12.5%)	27/195 (13.8%)	RR 0.9 (0.54 to 1.5)	14 fewer per 1000 (from 64 fewer to 69 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 12 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Kocsis 2009/ Klein 2011)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	2/200 (1%)	1/195 (0.51%)	RR 1.95 (0.18 to 21.33)	5 more per 1000 (from 4 fewer to 104 more)	LOW	CRITICAL
Functional impairment endpoint (follow-up mean 12 weeks; measured with: Longitudinal Interval Follow-up Evaluation Range of Impaired Functioning Tool (LIFE-RIFT); Better indicated by lower values)												
1 (Kocsis 2009/ Klein 2011)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	172	162	-	SMD 0.15 lower (0.36 lower to 0.07 higher)	HIGH	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] 95% CI crosses thresholds for both clinically important benefit and no effect
[2] 95% CI crosses threshold for no effect and thresholds for both clinically important benefit and harm

Table 72: Clinical evidence profile for comparison 3. Augmenting with counselling versus continuing with antidepressant

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with counselling**	**Continuing with antidepressant**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 12 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Kocsis 2009/ Klein 2011)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	168	76	-	SMD 0.06 higher (0.21 lower to 0.33 higher)	HIGH	CRITICAL
Remission (ITT) (follow-up mean 12 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D) AND responding (≥50% improvement on HAM-D))												
1 (Kocsis 2009/ Klein 2011)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	52/195 (26.7%)	30/96 (31.3%)	RR 0.85 (0.59 to 1.24)	47 fewer per 1000 (from 128 fewer to 75 more)	MODERATE	CRITICAL
Discontinuation due to any reason (follow-up mean 12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												

1 (Kocsis 2009/ Klein 2011)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	27/195 (13.8%)	16/96 (16.7%)	RR 0.83 (0.47 to 1.47)	28 fewer per 1000 (from 88 fewer to 78 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 12 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Kocsis 2009/ Klein 2011)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	1/195 (0.51%)	2/96 (2.1%)	RR 0.25 (0.02 to 2.68)	16 fewer per 1000 (from 20 fewer to 35 more)	LOW	CRITICAL
Functional impairment endpoint (follow-up mean 12 weeks; measured with: Longitudinal Interval Follow-up Evaluation Range of Impaired Functioning Tool (LIFE-RIFT); Better indicated by lower values)												
1 (Kocsis 2009/ Klein 2011)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	162	75	-	SMD 0.07 lower (0.34 lower to 0.21 higher)	HIGH	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] 95% CI crosses thresholds for both clinically important harm and no effect
[2] 95% CI crosses thresholds for no effect and both clinically important benefit and harm

Table 73: Clinical evidence profile for comparison 4. Augmenting with IPT versus continuing with antidepressant

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with IPT**	**Continuing with antidepressant**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up 5-19 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
2 (Schramm 2007, Souza 2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	79	79	-	SMD 0.36 lower (0.68 to 0.05 lower)	LOW	CRITICAL
Depression symptomatology change score (follow-up 5-19 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
3 (Murray 2010, Schramm 2007, Souza 2016)	randomised trials	no serious risk of bias	serious[3]	no serious indirectness	serious[2]	none	106	106	-	SMD 0.73 lower (1.38 to 0.08 lower)	LOW	CRITICAL
Depression symptomatology at 1-3 month follow-up (follow-up 1-3 months; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
2 (Schramm 2007, Souza 2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	66	65	-	SMD 0.31 lower (0.79 lower to 0.16 higher)	LOW	CRITICAL
Depression symptomatology at 12-month follow-up (follow-up mean 12 months; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												

1 (Schramm 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	50	47	-	SMD 0.54 lower (0.94 to 0.13 lower)	LOW	CRITICAL
Remission (ITT) (follow-up 5-19 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
4 (Murray 2010, Reynolds 2010, Schramm 2007, Souza 2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	83/176 (47.2%)	57/182 (31.3%)	RR 1.44 (1.12 to 1.86)	138 more per 1000 (from 38 more to 269 more)	LOW	CRITICAL
Response (ITT) (follow-up 5-19 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
3 (Murray 2010, Schramm 2007, Souza 2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	62/116 (53.4%)	40/118 (33.9%)	RR 1.51 (1.14 to 1.99)	173 more per 1000 (from 47 more to 336 more)	LOW	CRITICAL
Discontinuation due to any reason (follow-up 5-19 weeks; assessed with: Number of participants who dropped out for any reason)												
4 (Murray 2010, Reynolds 2010, Schramm 2007, Souza 2016)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	none	31/176 (17.6%)	23/182 (12.6%)	RR 1.35 (0.81 to 2.23)	44 more per 1000 (from 24 fewer to 155 more)	LOW	CRITICAL
Global functioning endpoint (follow-up mean 5 weeks; measured with: Global Assessment of Function (GAF); Better indicated by higher values)												
1 (Schramm 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	63	61	-	SMD 0.32 higher (0.03 lower to 0.68 higher)	LOW	IMPORTANT
Global functioning at 3-month follow-up (follow-up mean 3 months; measured with: Global Assessment of Function (GAF); Better indicated by higher values)												
1 (Schramm 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	50	47	-	SMD 0.44 higher (0.03 to 0.84 higher)	LOW	IMPORTANT
Global functioning at 12-month follow-up (follow-up mean 12 months; measured with: Global Assessment of Function (GAF); Better indicated by higher values)												
1 (Schramm 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	50	47	-	SMD 0.47 higher (0.06 to 0.87 higher)	LOW	IMPORTANT

CI: confidence interval; IPT: interpersonal therapy; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] Substantial heterogeneity
[4] 95% CI crosses thresholds for both clinically important harm and no effect

Table 74: Clinical evidence profile for comparison 5. Augmenting with short-term psychodynamic psychotherapy versus continuing with antidepressant

Quality assessment							No of patients		Effect	Quality	Importance

No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with short-term psychodynamic psychotherapy**	**Continuing with antidepressant**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 26 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Town 2017/2020)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	30	30	-	SMD 0.56 lower (1.07 to 0.04 lower)	MODERATE	CRITICAL
Depression symptomatology change score (follow-up mean 26 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Town 2017/2020)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	30	30	-	SMD 0.71 lower (1.23 to 0.19 lower)	MODERATE	CRITICAL
Depression symptomatology at 3-month follow-up (follow-up mean 3 months; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Town 2017/2020)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	30	30	-	SMD 0.58 lower (1.1 to 0.07 lower)	MODERATE	CRITICAL
Depression symptomatology at 6-month follow-up (follow-up mean 6 months; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Town 2017/2020)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	30	30	-	SMD 0.56 lower (1.08 to 0.05 lower)	MODERATE	CRITICAL
Depression symptomatology at 12-month follow-up (follow-up mean 12 months; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Town 2017/2020)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	30	30	-	SMD 0.62 lower (1.14 to 0.1 lower)	MODERATE	CRITICAL
Remission (ITT) (follow-up mean 26 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Town 2017/2020)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	11/30 (36.7%)	1/30 (3.3%)	RR 11 (1.51 to 79.96)	333 more per 1000 (from 17 more to 1000 more)	HIGH	CRITICAL
Remission (ITT) at 12-month follow-up (follow-up mean 12 months; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Town 2017/2020)	randomised trials	no serious	no serious inconsistency	no serious indirectness	very serious[2]	none	12/30 (40%)	9/30 (30%)	RR 1.33 (0.66 to 2.69)	99 more per 1000 (from 102	LOW	CRITICAL

		risk of bias								fewer to 507 more)		
Response (ITT) at 12-month follow-up (follow-up mean 12 months; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (Town 2017/2020)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	15/30 (50%)	12/30 (40%)	RR 1.25 (0.71 to 2.2)	100 more per 1000 (from 116 fewer to 480 more)	LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 26 weeks; assessed with: Number of participants who dropped out for any reason)												
1 (Town 2017/2020)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	5/30 (16.7%)	3/30 (10%)	RR 1.67 (0.44 to 6.36)	67 more per 1000 (from 56 fewer to 536 more)	LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] 95% CI crosses thresholds for both clinically important benefit and no effect
[2] 95% CI crosses thresholds for no effect and for both clinically important benefit and harm

Table 75: Clinical evidence profile for comparison 6. Augmenting with long-term psychodynamic psychotherapy versus continuing with antidepressant

Quality assessment							**No of patients**		**Effect**		**Quality**	**Importance**
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with long-term psychodynamic psychotherapy**	**Continuing with antidepressant**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 78 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Fonagy 2015)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	53	46	-	SMD 0.23 lower (0.63 lower to 0.16 higher)	VERY LOW	CRITICAL
Depression symptomatology at 6-month follow-up (follow-up mean 6 months; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Fonagy 2015)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	49	47	-	SMD 0.34 lower (0.75 lower to 0.06 higher)	VERY LOW	CRITICAL
Depression symptomatology at 12-month follow-up (follow-up mean 12 months; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Fonagy 2015)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	49	49	-	SMD 0.38 lower (0.78 lower to 0.02 higher)	VERY LOW	CRITICAL

Depression symptomatology at 24-month follow-up (follow-up mean 2 years; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Fonagy 2015)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	47	45	-	SMD 0.68 lower (1.1 to 0.26 lower)	VERY LOW	CRITICAL
Remission (ITT) (follow-up mean 78 weeks; assessed with: Number of people scoring <=8 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Fonagy 2015)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[3]	6/67 (9%)	4/62 (6.5%)	RR 1.39 (0.41 to 4.69)	25 more per 1000 (from 38 fewer to 238 more)	VERY LOW	CRITICAL
Remission (ITT) at 24-month follow-up (follow-up mean 2 years; assessed with: Number of people scoring <=8 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Fonagy 2015)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	10/67 (14.9%)	3/62 (4.8%)	RR 3.08 (0.89 to 10.69)	101 more per 1000 (from 5 fewer to 469 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 78 weeks; assessed with: Number of participants who dropped out for any reason)												
1 (Fornagy 2015)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[3]	10/67 (14.9%)	8/62 (12.9%)	RR 1.16 (0.49 to 2.74)	21 more per 1000 (from 66 fewer to 225 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Statistically significant group difference at baseline
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] Study partially funded by the International Psychoanalytic Association
[4] 95% CI crosses thresholds for no effect and for both clinically important benefit and harm

Table 76: Clinical evidence profile for comparison 7. Augmenting with self-help versus continuing with the antidepressant (+/- attention-placebo)

Quality assessment							**No of patients**		**Effect**		**Quality**	**Importance**
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with self-help**	**Continuing with the antidepressant (+/- attention-placebo)**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up 1.4-6 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) or Beck Depression Inventory (BDI-II); Better indicated by lower values)												
3 (Baert 2010_study 2, Dai 2019,	randomised trials	no serious	no serious inconsistency	no serious indirectness	serious[1]	none	80	77	-	SMD 0.29 lower (0.61	MODERATE	CRITICAL

Schlogelhofer 2014)		risk of bias								lower to 0.03 higher)		
Depression symptomatology change score (follow-up 1.4-6 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) or Beck Depression Inventory (BDI-II) change from baseline to endpoint; Better indicated by lower values)												
3 (Baert 2010_study 2, Dai 2019, Schlogelhofer 2014)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	80	77	-	SMD 0.39 lower (0.71 to 0.08 lower)	MODERATE	CRITICAL
Depression symptomatology at 1-month follow-up (follow-up mean 1 months; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Dai 2019)	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	16	16	-	SMD 1.37 lower (2.15 to 0.59 lower)	MODERATE	CRITICAL
Discontinuation due to any reason (follow-up 1.4-6 weeks; assessed with: Number of participants who dropped out for any reason)												
2 (Dai 2019, Schlogelhofer 2014)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	none	15/69 (21.7%)	10/61 (16.4%)	RR 1.32 (0.64 to 2.74)	52 more per 1000 (from 59 fewer to 285 more)	LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] 95% CI crosses thresholds for both clinically important benefit and no effect
[2] Risk of bias is high or unclear across multiple domains
[3] 95% CI crosses thresholds for no effect and for both clinically important benefit and harm

Table 77: Clinical evidence profile for comparison 8. Augmenting with self-help and switching to SSRI versus switching to SSRI-only

Quality assessment							**No of patients**		**Effect**		**Quality**	**Importance**
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with self-help and switching to SSRI**	**Switching to SSRI-only**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 9 weeks; measured with: Patient Health Questionnaire (PHQ-9); Better indicated by lower values)												
1 (Mantani 2017)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[2]	81	83	-	SMD 1.13 lower (1.46 to 0.8 lower)	LOW	CRITICAL
Depression symptomatology change score (follow-up mean 9 weeks; measured with: Patient Health Questionnaire (PHQ-9) change from baseline to endpoint; Better indicated by lower values)												
1 (Mantani 2017)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[2]	81	83	-	SMD 0.76 lower (1.08 to 0.44 lower)	VERY LOW	CRITICAL
Remission (ITT) (follow-up mean 9 weeks; assessed with: Number of people scoring <=4 on Patient Health Questionnaire (PHQ-9))												

1 (Mantani 2017)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[2]	25/81 (30.9%)	15/83 (18.1%)	RR 1.71 (0.97 to 3)	128 more per 1000 (from 5 fewer to 361 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 9 weeks; assessed with: Number of people showing at least 50% improvement on Patient Health Questionnaire (PHQ-9))												
1 (Mantani 2017)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[2]	34/81 (42%)	18/83 (21.7%)	RR 1.94 (1.19 to 3.14)	204 more per 1000 (from 41 more to 464 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 9 weeks; assessed with: Number of participants who dropped out for any reason)												
1 (Mantani 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[2]	1/81 (1.2%)	0/83 (0%)	RR 3.07 (0.13 to 74.35)	-	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SSRI: selective serotonin reuptake inhibitor
[1] Risk of bias is high or unclear across multiple domains
[2] Study partially funded by pharmaceutical companies
[3] 95% CI crosses thresholds for both clinically important benefit and no effect
[4] 95% CI crosses thresholds for no effect and both clinically important benefit and harm

Table 78: Clinical evidence profile for comparison 9. Augmenting with art therapy versus attention-placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with art therapy**	**Attention-placebo**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 6 weeks; measured with: Beck Depression Inventory (BDI-II); Better indicated by lower values)												
1 (Nan 2017)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	52	48	-	SMD 0.56 lower (0.96 to 0.16 lower)	LOW	CRITICAL
Depression symptomatology change score (follow-up mean 6 weeks; measured with: Beck Depression Inventory (BDI-II) change from baseline to endpoint; Better indicated by lower values)												
1 (Nan 2017)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	52	48	-	SMD 1.22 lower (1.64 to 0.79 lower)	MODERATE	CRITICAL
Discontinuation due to any reason (follow-up mean 6 weeks; assessed with: Number of participants who dropped out for any reason)												
1 (Nan 2017)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	none	1/53 (1.9%)	5/53 (9.4%)	RR 0.2 (0.02 to 1.65)	75 fewer per 1000 (from 92 fewer to 61 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference

[1] *Risk of bias is high or unclear across multiple domains*
[2] *95% CI crosses thresholds for both clinically important benefit and no effect*
[3] *95% CI crosses thresholds for no effect and both clinically important benefit and harm*

Table 79: Clinical evidence profile for comparison 10. Augmenting with eye movement desensitization reprocessing (EMDR) versus augmenting with cognitive behavioural therapy

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Augmenting with eye movement desensitization reprocessing (EMDR)	Augmenting with cognitive behavioural therapy	Relative (95% CI)	Absolute		
Depression symptomatology endpoint (follow-up 13-26 weeks; measured with: Beck Depression Inventory (BDI-II); Better indicated by lower values)												
1 (Ostacoli 2018)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	31	35	-	SMD 0.65 lower (1.14 to 0.15 lower)	VERY LOW	CRITICAL
Remission (ITT) (follow-up 13-26 weeks; assessed with: Number of people scoring <13 on Beck Depression Inventory (BDI-II))												
1 (Ostacoli 2018)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	22/40 (55%)	17/42 (40.5%)	RR 1.36 (0.86 to 2.16)	146 more per 1000 (from 57 fewer to 470 more)	VERY LOW	CRITICAL
Remission (ITT) at 6-month follow-up (follow-up mean 6 months; assessed with: Number of people scoring <13 on Beck Depression Inventory (BDI-II))												
1 (Ostacoli 2018)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[3]	17/40 (42.5%)	15/42 (35.7%)	RR 1.19 (0.69 to 2.05)	68 more per 1000 (from 111 fewer to 375 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up 13-26 weeks; assessed with: Number of participants who dropped out for any reason)												
1 (Ostacoli 2018)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[3]	9/40 (22.5%)	7/42 (16.7%)	RR 1.35 (0.56 to 3.28)	58 more per 1000 (from 73 fewer to 380 more)	VERY LOW	CRITICAL
Global functioning at endpoint (follow-up 13-26 weeks; measured with: Global Assessment of Function (GAF); Better indicated by higher values)												
1 (Ostacoli 2018)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	31	35	-	SMD 0.22 higher (0.27 lower to 0.7 higher)	VERY LOW	IMPORTANT
Global functioning at 6-month follow-up (follow-up mean 6 months; measured with: Global Assessment of Function (GAF); Better indicated by higher values)												

1 (Ostacoli 2018)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	31	35	-	SMD 0.24 higher (0.24 lower to 0.73 higher)	VERY LOW	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias high or unclear across multiple domains
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] Potential conflict of interest as study funded by the EMDR Research Foundation
[4] 95% CI crosses thresholds for no effect and both clinically important benefit and harm

Table 80: Clinical evidence profile for comparison 11. Increasing the dose of SSRI versus continuing SSRI at the same dose

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Increasing the dose of SSRI**	**Continuing SSRI at the same dose**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 6 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Ruhe 2009)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	30	27	-	SMD 0.63 higher (0.1 to 1.17 higher)	MODERATE	CRITICAL
Depression symptomatology change score (follow-up 5-6 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) or Montgomery Asberg Depression Rating Scale (MADRS) change from baseline to endpoint; Better indicated by lower values)												
2 (Dornseif 1989, Kim 2019)	randomised trials	serious[2]	serious[3]	no serious indirectness	serious[4]	reporting bias[5]	205	211	-	SMD 0.33 lower (0.73 lower to 0.07 higher)	VERY LOW	CRITICAL
Remission (ITT) (follow-up 5-6 weeks; assessed with: Number of people scoring <=7/<=8 on Hamilton Rating Scale for Depression (HAM-D) or <=10 on Montgomery Asberg Depression Rating Scale (MADRS))												
5 (Dornseif 1989, Kim 2019, Licht 2002, Ruhe 2009, Schweizer 2001)	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[4]	reporting bias[5]	116/372 (31.2%)	112/381 (29.4%)	RR 1.1 (0.84 to 1.45)	29 more per 1000 (from 47 fewer to 132 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up 5-6 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D)/Montgomery Asberg Depression Rating Scale (MADRS) or rated as much or very much improved on Clinical Global Impressions scale (CGI-I))												
6 (Dornseif 1989, Kim 2019, Licht 2002, Ruhe 2009, Schweizer	randomised trials	serious[2]	serious[3]	no serious indirectness	serious[4]	reporting bias[5]	195/408 (47.8%)	195/422 (46.2%)	RR 1.1 (0.86 to 1.39)	46 more per 1000 (from 65	VERY LOW	CRITICAL

1990, Schweizer 2001)										fewer to 180 more)		
Discontinuation due to any reason (follow-up 5-6 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
5 (Dornseif 1989, Kim 2019, Licht 2002, Ruhe 2009, Schweizer 2001)	randomised trials	no serious risk of bias	serious[3]	no serious indirectness	very serious[6]	reporting bias[5]	66/372 (17.7%)	77/381 (20.2%)	RR 0.77 (0.4 to 1.48)	46 fewer per 1000 (from 121 fewer to 97 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up 5-6 weeks; assessed with: Number of participants who dropped out due to adverse events)												
4 (Dornseif 1989, Kim 2019, Ruhe 2009, Schweizer 1990)	randomised trials	no serious risk of bias	serious[3]	no serious indirectness	very serious[6]	reporting bias[5]	27/272 (9.9%)	16/286 (5.6%)	RR 1.59 (0.42 to 6.03)	33 more per 1000 (from 32 fewer to 281 more)	VERY LOW	CRITICAL
Quality of life physical component score (PCS) endpoint (follow-up mean 6 weeks; measured with: 36-item Short-Form Survey (SF-36): Physical component score; Better indicated by higher values)												
1 (Ruhe 2009)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	30	27	-	SMD 0.6 lower (1.13 to 0.06 lower)	MODERATE	IMPORTANT
Quality of life mental component score (MCS) endpoint (follow-up mean 6 weeks; measured with: 36-item Short-Form Survey (SF-36): Mental component score; Better indicated by higher values)												
1 (Ruhe 2009)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	30	27	-	SMD 1.55 higher (0.95 to 2.14 higher)	HIGH	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SSRI: selective serotonin reuptake inhibitor
[1] 95% CI crosses thresholds for both clinically important harm and no effect
[2] Risk of bias is high or unclear across multiple domains
[3] Substantial heterogeneity
[4] 95% CI crosses thresholds for both clinically important benefit and no effect
[5] Funding from pharmaceutical companies
[6] 95% CI crosses thresholds for no effect and both clinically important benefit and harm

Table 81: Clinical evidence profile for comparison 12. Increasing the dose of SSRI versus switching to SNRI

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Increasing the dose of SSRI**	**Switching to SNRI**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 8 weeks; measured with: Quick Inventory of Depressive Symptomatology (QIDS); Better indicated by lower values)												

1 (Bose 2012)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[2]	229	243	-	SMD 0.21 lower (0.39 to 0.03 lower)	LOW	CRITICAL
Depression symptomatology change score (follow-up mean 8 weeks; measured with: Quick Inventory of Depressive Symptomatology (QIDS) change from baseline to endpoint; Better indicated by lower values)												
1 (Bose 2012)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[2]	229	243	-	SMD 0.16 lower (0.35 lower to 0.02 higher)	LOW	CRITICAL
Remission (ITT) (follow-up mean 8 weeks; assessed with: Number of people scoring <=10 on Montgomery Asberg Depression Rating Scale (MADRS))												
1 (Bose 2012)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[2]	124/238 (52.1%)	102/246 (41.5%)	RR 1.26 (1.04 to 1.52)	108 more per 1000 (from 17 more to 216 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 8 weeks; assessed with: Number of people showing at least 50% improvement on Montgomery Asberg Depression Rating Scale (MADRS))												
1 (Bose 2012)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[2]	167/238 (70.2%)	170/246 (69.1%)	RR 1.02 (0.9 to 1.14)	14 more per 1000 (from 69 fewer to 97 more)	LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 8 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Bose 2012)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[2]	56/238 (23.5%)	53/246 (21.5%)	RR 1.09 (0.78 to 1.52)	19 more per 1000 (from 47 fewer to 112 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 8 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Bose 2012)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[2]	13/238 (5.5%)	13/246 (5.3%)	RR 1.03 (0.49 to 2.18)	2 more per 1000 (from 27 fewer to 62 more)	VERY LOW	CRITICAL
Quality of life endpoint (follow-up mean 8 weeks; measured with: Quality of Life Enjoyment and Satisfaction Questionnaire-short form (Q-LES-Q-SF); Better indicated by higher values)												
1 (Bose 2012)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[2]	229	243	-	SMD 0.11 higher (0.08 lower to 0.29 higher)	LOW	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SNRI: serotonin-norepinephrine reuptake inhibitor; SSRI: selective serotonin reuptake inhibitor

[1] Risk of bias is high or unclear across multiple domains

[2] Funding from pharmaceutical company

[3] 95% CI crosses thresholds for both clinically important benefit and no effect

[4] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 82: Clinical evidence profile for comparison 13. Increasing the dose of SSRI versus augmenting with TCA

Quality assessment							No of patients		Effect	Quality	Importance

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Increasing the dose of SSRI	Augmenting with TCA	Relative (95% CI)	Absolute		
Depression symptomatology endpoint (follow-up mean 4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
2 (Fava 1994a, Fava 2002)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	48	46	-	SMD 0.67 lower (1.28 to 0.05 lower)	LOW	CRITICAL
Depression symptomatology change score (follow-up mean 4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
2 (Fava 1994a, Fava 2002)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	48	46	-	SMD 0.44 lower (0.9 lower to 0.01 higher)	LOW	CRITICAL
Remission (ITT) (follow-up mean 4 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
2 (Fava 1994a, Fava 2002)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	22/48 (45.8%)	13/46 (28.3%)	RR 1.6 (0.91 to 2.81)	170 more per 1000 (from 25 fewer to 512 more)	LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 4 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
2 (Fava 1994a, Fava 2002)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	none	5/48 (10.4%)	8/46 (17.4%)	RR 0.58 (0.21 to 1.64)	73 fewer per 1000 (from 137 fewer to 111 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 4 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Fava 1994a)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	reporting bias[4]	0/15 (0%)	2/12 (16.7%)	RR 0.16 (0.01 to 3.09)	140 fewer per 1000 (from 165 fewer to 348 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SSRI: selective serotonin reuptake inhibitor; TCA: tricyclic antidepressant
[1] Risk of bias is high or unclear across multiple domains
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[4] Study partially funded by pharmaceutical company

Table 83: Clinical evidence profile for comparison 14. Increasing the dose of SSRI versus augmenting with antipsychotic

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Increasing the dose of SSRI**	**Augmenting with antipsychotic**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 13 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												

1 (Rocca 2002b)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	28	32	-	SMD 0.1 higher (0.41 lower to 0.6 higher)	MODERATE	CRITICAL
Depression symptomatology change score (follow-up mean 13 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Rocca 2002b)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	28	32	-	SMD 0.07 higher (0.43 lower to 0.58 higher)	MODERATE	CRITICAL
Remission (ITT) (follow-up mean 13 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Rocca 2002b)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	9/28 (32.1%)	14/32 (43.8%)	RR 0.73 (0.38 to 1.43)	118 fewer per 1000 (from 271 fewer to 188 more)	LOW	CRITICAL
Response (ITT) (follow-up mean 13 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (Rocca 2002b)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	15/28 (53.6%)	18/32 (56.3%)	RR 0.95 (0.6 to 1.51)	28 fewer per 1000 (from 225 fewer to 287 more)	LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 13 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Rocca 2002b)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	4/28 (14.3%)	5/32 (15.6%)	RR 0.91 (0.27 to 3.08)	14 fewer per 1000 (from 114 fewer to 325 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 13 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Rocca 2002b)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	2/28 (7.1%)	2/32 (6.3%)	RR 1.14 (0.17 to 7.59)	9 more per 1000 (from 52 fewer to 412 more)	LOW	CRITICAL
Functional remission (follow-up mean 13 weeks; assessed with: Number of people scoring =>71 on Global Assessment of Function (GAF))												
1 (Rocca 2002b)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	11/28 (39.3%)	22/32 (68.8%)	RR 0.57 (0.34 to 0.96)	296 fewer per 1000 (from 28 fewer to 454 fewer)	MODERATE	IMPORTANT
Global functioning endpoint (follow-up mean 13 weeks; measured with: Global Assessment of Function (GAF); Better indicated by higher values)												
1 (Rocca 2002b)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	28	32	-	SMD 0.67 lower (1.19 to 0.15 lower)	MODERATE	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SSRI: selective serotonin reuptake inhibitor
[1] 95% CI crosses thresholds for both clinically important harm and no effect
[2] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 84: Clinical evidence profile for comparison 15. Increasing the dose of SSRI versus augmenting with lithium

Quality assessment							No of patients		Effect	Quality	Importance

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Increasing the dose of SSRI	Augmenting with lithium	Relative (95% CI)	Absolute		
Depression symptomatology endpoint (follow-up mean 4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
2 (Fava 1994a, Fava 2002)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	48	48	-	SMD 0.34 lower (0.75 lower to 0.06 higher)	LOW	CRITICAL
Depression symptomatology change score (follow-up mean 4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
2 (Fava 1994a, Fava 2002)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	48	48	-	SMD 0.31 lower (0.72 lower to 0.09 higher)	LOW	CRITICAL
Remission (ITT) (follow-up mean 4 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
2 (Fava 1994a, Fava 2002)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	22/48 (45.8%)	12/48 (25%)	RR 1.83 (1.03 to 3.25)	208 more per 1000 (from 7 more to 562 more)	LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 4 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
2 (Fava 1994a, Fava 2002)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	none	5/48 (10.4%)	7/48 (14.6%)	RR 0.72 (0.24 to 2.11)	41 fewer per 1000 (from 111 fewer to 162 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 4 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Fava 1994a)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	reporting bias[4]	0/15 (0%)	1/14 (7.1%)	RR 0.31 (0.01 to 7.09)	49 fewer per 1000 (from 71 fewer to 435 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SSRI: selective serotonin reuptake inhibitor
[1] Risk of bias was high or unclear across multiple domains
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[4] Study partially funded by pharmaceutical company

Table 85: Clinical evidence profile for comparison 16. Switching to SSRI versus continuing with antidepressant

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Switching to SSRI**	**Continuing with antidepressant**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology change score (follow-up 8-12 weeks; measured with: Montgomery Asberg Depression Rating Scale (MADRS) change from baseline to endpoint; Better indicated by lower values)												

2 (Corya 2006, Shelton 2005)	randomised trials	serious[1]	serious[2]	no serious indirectness	no serious imprecision	reporting bias[3]	198	126	-	SMD 0.03 higher (0.31 lower to 0.38 higher)	VERY LOW	CRITICAL
Remission (ITT) (follow-up 8-12 weeks; assessed with: Number of people scoring <=8 on Montgomery Asberg Depression Rating Scale (MADRS))												
2 (Corya 2006, Shelton 2005)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	reporting bias[3]	29/202 (14.4%)	25/127 (19.7%)	RR 0.76 (0.46 to 1.24)	47 fewer per 1000 (from 106 fewer to 47 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up 8-12 weeks; assessed with: Number of people showing at least 50% improvement on Montgomery Asberg Depression Rating Scale (MADRS))												
2 (Corya 2006, Shelton 2005)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	reporting bias[3]	60/202 (29.7%)	50/127 (39.4%)	RR 0.78 (0.54 to 1.12)	87 fewer per 1000 (from 181 fewer to 47 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up 8-12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
2 (Corya 2006, Shelton 2005)	randomised trials	no serious risk of bias	serious[2]	no serious indirectness	very serious[5]	reporting bias[3]	40/202 (19.8%)	23/127 (18.1%)	RR 1.13 (0.54 to 2.38)	24 more per 1000 (from 83 fewer to 250 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up 8-12 weeks; assessed with: Number of participants who dropped out due to adverse events)												
2 (Corya 2006, Shelton 2005)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[5]	reporting bias[3]	7/202 (3.5%)	3/127 (2.4%)	RR 1.43 (0.38 to 5.47)	10 more per 1000 (from 15 fewer to 106 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SSRI: selective serotonin reuptake inhibitor
[1] Risk of bias is high or unclear across multiple domains
[2] Substantial heterogeneity
[3] Funding from pharmaceutical companies
[4] 95% CI crosses thresholds for both clinically important harm and no effect
[5] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 86: Clinical evidence profile for comparison 17. Switching to a different SSRI versus continuing same SSRI

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Switching to a different SSRI**	**Continuing same SSRI**	**Relative (95% CI)**	**Absolute**		
Remission (ITT) (follow-up mean 6 weeks; assessed with: Number of people scoring <=10 on Montgomery Asberg Depression Rating Scale (MADRS))												

1 (Nakajima 2011)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[2]	12/20 (60%)	3/21 (14.3%)	RR 4.2 (1.39 to 12.71)	457 more per 1000 (from 56 more to 1000 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 6 weeks; assessed with: Number of people showing at least 50% improvement on Montgomery Asberg Depression Rating Scale (MADRS))												
1 (Nakajima 2011)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[2]	15/20 (75%)	4/21 (19%)	RR 3.94 (1.57 to 9.85)	560 more per 1000 (from 109 more to 1000 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 6 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Nakajima 2011)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	reporting bias[2]	2/20 (10%)	5/21 (23.8%)	RR 0.42 (0.09 to 1.92)	138 fewer per 1000 (from 217 fewer to 219 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 6 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Nakajima 2011)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[2]	0/20 (0%)	0/21 (0%)	not pooled	not pooled	LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SSRI: selective serotonin reuptake inhibitor
[1] Risk of bias is high or unclear across multiple domains and abrupt tapering of failed drug in switch arm
[2] Study partially funded by pharmaceutical company
[3] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 87: Clinical evidence profile for comparison 18. Switching to SSRI versus antipsychotic

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Switching to SSRI**	**Antipsychotic**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology change score (follow-up 8-12 weeks; measured with: Montgomery Asberg Depression Rating Scale (MADRS) change from baseline to endpoint; Better indicated by lower values)												
2 (Corya 2006, Shelton 2005)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	198	203	-	SMD 0.27 lower (0.5 to 0.03 lower)	VERY LOW	CRITICAL
Remission (ITT) (follow-up 8-12 weeks; assessed with: Number of people scoring <=8 on Montgomery Asberg Depression Rating Scale (MADRS))												
2 (Corya 2006, Shelton 2005)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[3]	29/202 (14.4%)	27/206 (13.1%)	RR 1.1 (0.67 to 1.79)	13 more per 1000 (from 43 fewer to 104 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up 8-12 weeks; assessed with: Number of people showing at least 50% improvement on Montgomery Asberg Depression Rating Scale (MADRS))												
2 (Corya 2006, Shelton 2005)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	60/202 (29.7%)	43/206 (20.9%)	RR 1.42 (1.01 to 2)	88 more per 1000 (from 2 more to 209 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up 8-12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												

2 (Corya 2006, Shelton 2005)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	40/202 (19.8%)	50/206 (24.3%)	RR 0.82 (0.56 to 1.18)	44 fewer per 1000 (from 107 fewer to 44 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up 8-12 weeks; assessed with: Number of participants who dropped out due to adverse events)												
2 (Corya 2006, Shelton 2005)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	7/202 (3.5%)	19/206 (9.2%)	RR 0.39 (0.16 to 0.91)	56 fewer per 1000 (from 8 fewer to 77 fewer)	LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SSRI: selective serotonin reuptake inhibitor
[1] Risk of bias is high or unclear across multiple domains
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] Funding from pharmaceutical companies
[4] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 88: Clinical evidence profile for comparison 19. Switching to combined SSRI + antipsychotic versus switching to antipsychotic-only

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Switching to combined SSRI + antipsychotic**	**Switching to antipsychotic-only**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology change score (follow-up 8-12 weeks; measured with: Montgomery Asberg Depression Rating Scale (MADRS) change from baseline to endpoint; Better indicated by lower values)												
2 (Corya 2006, Shelton 2005)	randomised trials	serious[1]	very serious[2]	no serious indirectness	serious[3]	reporting bias[4]	376	203	-	SMD 0.44 lower (0.91 lower to 0.03 higher)	VERY LOW	CRITICAL
Remission (ITT) (follow-up 8-12 weeks; assessed with: Number of people scoring <=8 on Montgomery Asberg Depression Rating Scale (MADRS))												
2 (Corya 2006, Shelton 2005)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[4]	94/389 (24.2%)	27/206 (13.1%)	RR 1.63 (0.97 to 2.73)	83 more per 1000 (from 4 fewer to 227 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up 8-12 weeks; assessed with: Number of people showing at least 50% improvement on Montgomery Asberg Depression Rating Scale (MADRS))												
2 (Corya 2006, Shelton 2005)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[4]	140/389 (36%)	43/206 (20.9%)	RR 1.53 (1.12 to 2.1)	111 more per 1000 (from 25 more to 230 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up 8-12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												

2 (Corya 2006, Shelton 2005)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[4]	90/389 (23.1%)	50/206 (24.3%)	RR 0.89 (0.65 to 1.21)	27 fewer per 1000 (from 85 fewer to 51 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up 8-12 weeks; assessed with: Number of participants who dropped out due to adverse events)												
2 (Corya 2006, Shelton 2005)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[5]	reporting bias[4]	39/389 (10%)	19/206 (9.2%)	RR 0.98 (0.48 to 2.03)	2 fewer per 1000 (from 48 fewer to 95 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SSRI: selective serotonin reuptake inhibitor
[1] Risk of bias is high or unclear across multiple domains
[2] Considerable heterogeneity
[3] 95% CI crosses thresholds for both clinically important benefit and no effect
[4] Funding from pharmaceutical companies
[5] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 89: Clinical evidence profile for comparison 20. Augmenting with SSRI versus augmenting with lithium

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with SSRI**	**Augmenting with lithium**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology change score (follow-up mean 10 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Navarro 2019b)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	52	52	-	SMD 0.56 lower (0.95 to 0.16 lower)	LOW	CRITICAL
Remission (ITT) (follow-up mean 10 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Navarro 2019b)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	21/52 (40.4%)	11/52 (21.2%)	RR 1.91 (1.03 to 3.55)	193 more per 1000 (from 6 more to 539 more)	LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SSRI: selective serotonin reuptake inhibitor
[1] Risk of bias is high or unclear across multiple domains
[2] 95% CI crosses thresholds for both clinically important benefit and no effect

Table 90: Clinical evidence profile for comparison 21. Switching to TCA versus SSRI

Quality assessment	No of patients	Effect	Quality	Importance

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Switching to TCA	SSRI	Relative (95% CI)	Absolute		
Depression symptomatology endpoint (follow-up mean 4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Souery 2011a)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	67	85	-	SMD 0.2 lower (0.52 lower to 0.12 higher)	VERY LOW	CRITICAL
Remission (ITT) (follow-up mean 4 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Souery 2011a)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	21/84 (25%)	16/105 (15.2%)	RR 1.64 (0.92 to 2.94)	98 more per 1000 (from 12 fewer to 296 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 4 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (Souery 2011a)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[3]	37/84 (44%)	46/105 (43.8%)	RR 1.01 (0.73 to 1.39)	4 more per 1000 (from 118 fewer to 171 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 4 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Souery 2011a)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[3]	17/84 (20.2%)	20/105 (19%)	RR 1.06 (0.6 to 1.9)	11 more per 1000 (from 76 fewer to 171 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SSRI: selective serotonin reuptake inhibitor; TCA: tricyclic antidepressant
[1] Risk of bias is high or unclear across multiple domains
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] Study partially funded by pharmaceutical company
[4] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 91: Clinical evidence profile for comparison 22. Switching to TCA versus augmenting with mirtazapine

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Switching to TCA**	**Augmenting with mirtazapine**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 10 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Navarro 2019a)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	56	56	-	SMD 1.13 lower (1.53 to 0.73 lower)	LOW	CRITICAL
Depression symptomatology change score (follow-up mean 10 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Navarro 2019a)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	56	56	-	SMD 1.47 lower (1.88 to 1.05 lower)	LOW	CRITICAL
Remission (ITT) (follow-up mean 10 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												

1 (Navarro 2019a)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	40/56 (71.4%)	22/56 (39.3%)	RR 1.82 (1.26 to 2.62)	322 more per 1000 (from 102 more to 636 more)	LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 10 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Navarro 2019a)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	5/56 (8.9%)	2/56 (3.6%)	RR 2.5 (0.51 to 12.35)	54 more per 1000 (from 18 fewer to 405 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; TCA: tricyclic antidepressant
[1] Risk of bias is high or unclear across multiple domains and rapid tapering of failed drug in switch arm
[2] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 92: Clinical evidence profile for comparison 23. Switching to mianserin versus continuing with antidepressant

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Switching to mianserin**	**Continuing with antidepressant**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology change score (follow-up mean 6 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Ferreri 2001)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	33	38	-	SMD 0.24 lower (0.71 lower to 0.23 higher)	VERY LOW	CRITICAL
Remission (ITT) (follow-up mean 6 weeks; assessed with: Number of people scoring <=8 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Ferreri 2001)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	12/34 (35.3%)	7/38 (18.4%)	RR 1.92 (0.85 to 4.3)	169 more per 1000 (from 28 fewer to 608 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 6 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (Ferreri 2001)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[3]	16/34 (47.1%)	14/38 (36.8%)	RR 1.28 (0.74 to 2.21)	103 more per 1000 (from 96 fewer to 446 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 6 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Ferreri 2001)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[5]	reporting bias[3]	12/34 (35.3%)	7/38 (18.4%)	RR 1.92 (0.85 to 4.3)	169 more per 1000 (from 28 fewer to 608 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 6 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Ferreri 2001)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[5]	reporting bias[3]	8/34 (23.5%)	0/38 (0%)	RR 18.94 (1.13 to 316.35)	-	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference

[1] *Risk of bias is high or unclear across multiple domains and abrupt tapering for the switch arm*
[2] *95% CI crosses thresholds for both clinically important benefit and no effect*
[3] *Study funded by pharmaceutical company*
[4] *95% CI crosses thresholds for no effect, and both clinically important benefit and harm*
[5] *95% CI crosses thresholds for both clinically important harm and no effect*

Table 93: Clinical evidence profile for comparison 24. Augmenting with mianserin versus continuing with antidepressant (+/- placebo)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Augmenting with mianserin	Continuing with antidepressant (+/- placebo)	Relative (95% CI)	Absolute		
Depression symptomatology change score (follow-up mean 6 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Ferreri 2001)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	32	38	-	SMD 0.66 lower (1.14 to 0.17 lower)	VERY LOW	CRITICAL
Remission (ITT) (follow-up 5-6 weeks; assessed with: Number of people scoring <=7/<=8 on Hamilton Rating Scale for Depression (HAM-D))												
2 (Ferreri 2001, Licht 2002)	randomised trials	serious[1]	serious[4]	no serious indirectness	very serious[5]	reporting bias[3]	57/130 (43.8%)	44/137 (32.1%)	RR 1.53 (0.78 to 2.99)	170 more per 1000 (from 71 fewer to 639 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up 5-6 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
2 (Ferreri 2001, Licht 2002)	randomised trials	serious[1]	serious[4]	no serious indirectness	very serious[5]	reporting bias[3]	86/130 (66.2%)	83/137 (60.6%)	RR 1.22 (0.7 to 2.13)	133 more per 1000 (from 182 fewer to 685 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up 5-6 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
2 (Ferreri 2001, Licht 2002)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[5]	reporting bias[3]	23/130 (17.7%)	17/137 (12.4%)	RR 1.43 (0.79 to 2.56)	53 more per 1000 (from 26 fewer to 194 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 6 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Ferreri 2001)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[5]	reporting bias[3]	2/32 (6.3%)	0/38 (0%)	RR 5.91 (0.29 to 118.78)	-	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] *Risk of bias is high or unclear across multiple domains*
[2] *95% CI crosses thresholds for both clinically important benefit and harm*

[3] *Funding from pharmaceutical company*
[4] *Substantial heterogeneity*
[5] *95% CI crosses thresholds for no effect, and both clinically important benefit and harm*

Table 94: Clinical evidence profile for comparison 25. Augmenting with mianserin versus increasing dose of antidepressant

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Augmenting with mianserin	Increasing dose of antidepressant	Relative (95% CI)	Absolute		
Remission (ITT) (follow-up mean 5 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Licht 2002)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	43/98 (43.9%)	28/98 (28.6%)	RR 1.54 (1.05 to 2.26)	154 more per 1000 (from 14 more to 360 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 5 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (Licht 2002)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	66/98 (67.3%)	54/98 (55.1%)	RR 1.22 (0.98 to 1.53)	121 more per 1000 (from 11 fewer to 292 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 5 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Licht 2002)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[3]	17/98 (17.3%)	15/98 (15.3%)	RR 1.13 (0.6 to 2.14)	20 more per 1000 (from 61 fewer to 174 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk
[1] *Risk of bias is high or unclear across multiple domains*
[2] *95% CI crosses thresholds for both clinically important benefit and no effect*
[3] *Study funded by pharmaceutical company*
[4] *95% CI crosses thresholds for no effect, and both clinically important benefit and harm*

Table 95: Clinical evidence profile for comparison 26. Augmenting with mianserin versus switch to mianserin

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Augmenting with mianserin	Switch to mianserin	Relative (95% CI)	Absolute		
Depression symptomatology change score (follow-up mean 6 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												

1 (Ferreri 2001)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	32	33	-	SMD 0.41 lower (0.91 lower to 0.08 higher)	VERY LOW	CRITICAL
Remission (ITT) (follow-up mean 6 weeks; assessed with: Number of people scoring <=8 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Ferreri 2001)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[3]	14/32 (43.8%)	12/34 (35.3%)	RR 1.24 (0.68 to 2.26)	85 more per 1000 (from 113 fewer to 445 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 6 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (Ferreri 2001)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	20/32 (62.5%)	16/34 (47.1%)	RR 1.33 (0.85 to 2.08)	155 more per 1000 (from 71 fewer to 508 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 6 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Ferreri 2001)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[3]	6/32 (18.8%)	12/34 (35.3%)	RR 0.53 (0.23 to 1.25)	166 fewer per 1000 (from 272 fewer to 88 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 6 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Ferreri 2001)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	2/32 (6.3%)	8/34 (23.5%)	RR 0.27 (0.06 to 1.16)	172 fewer per 1000 (from 221 fewer to 38 more)	LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains and abrupt tapering for the switch arm
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] Study funded by pharmaceutical company
[4] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 96: Clinical evidence profile for comparison 27. Increasing the dose of SNRI versus continuing SNRI at the same dose

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Increasing the dose of SNRI**	**Continuing SNRI at the same dose**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology change score (follow-up mean 8 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Kornstein 2008)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[2]	118	130	-	SMD 0.01 higher (0.24 lower to 0.26 higher)	VERY LOW	CRITICAL
Remission (ITT) (follow-up mean 8 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Kornstein 2008)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	reporting bias[2]	36/124 (29%)	39/131 (29.8%)	RR 0.98 (0.67 to 1.43)	6 fewer per 1000 (from 98 fewer to 128 more)	VERY LOW	CRITICAL

Response (ITT) (follow-up mean 8 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (Kornstein 2008)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	serious[4]	reporting bias[2]	48/124 (38.7%)	58/131 (44.3%)	RR 0.87 (0.65 to 1.17)	58 fewer per 1000 (from 155 fewer to 75 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 8 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Kornstein 2008)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	reporting bias[2]	34/124 (27.4%)	26/131 (19.8%)	RR 1.38 (0.88 to 2.16)	75 more per 1000 (from 24 fewer to 230 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 8 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Kornstein 2008)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	reporting bias[2]	7/124 (5.6%)	6/131 (4.6%)	RR 1.23 (0.43 to 3.57)	11 more per 1000 (from 26 fewer to 118 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SNRI: serotonin-norepinephrine reuptake inhibitor
[1] Risk of bias is high or unclear across multiple domains
[2] Study funded by pharmaceutical company
[3] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[4] 95% CI crosses thresholds for both clinically important harm and no effect

Table 97: Clinical evidence profile for comparison 28. Switching to SNRI versus continuing with antidepressant

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Switching to SNRI**	**Continuing with antidepressant**	**Relative (95% CI)**	**Absolute**		
Remission (ITT) (follow-up mean 8 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Fang 2010)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	21/50 (42%)	21/45 (46.7%)	RR 0.9 (0.57 to 1.41)	47 fewer per 1000 (from 201 fewer to 191 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 8 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (Fang 2010)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	32/50 (64%)	30/45 (66.7%)	RR 0.96 (0.72 to 1.29)	27 fewer per 1000 (from 187 fewer to 193 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 8 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Fang 2010)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	9/50 (18%)	8/45 (17.8%)	RR 1.01 (0.43 to 2.4)	2 more per 1000 (from 101 fewer to 249 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 8 weeks; assessed with: Number of participants who dropped out due to adverse events)												

1 (Fang 2010)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	0/50 (0%)	1/45 (2.2%)	RR 0.3 (0.01 to 7.2)	16 fewer per 1000 (from 22 fewer to 138 more)	VERY LOW	CRITICAL
Quality of life physical component score (PCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Physical component score; Better indicated by higher values)												
1 (Fang 2010)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	50	45	-	SMD 0.02 higher (0.38 lower to 0.42 higher)	LOW	IMPORTANT
Quality of life mental component score (MCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Mental component score; Better indicated by higher values)												
1 (Fang 2010)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	50	45	-	SMD 0.14 higher (0.26 lower to 0.54 higher)	VERY LOW	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SNRI: serotonin-norepinephrine reuptake inhibitor
[1] Risk of bias is high or unclear across multiple domains
[2] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[3] 95% CI crosses thresholds for both clinically important benefit and no effect

Table 98: Clinical evidence profile for comparison 29. Switching to SNRI versus switching to another antidepressant from same class

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Switching to SNRI**	**Switching to another antidepressant from same class**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology change score (follow-up 4-14 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) or Quick Inventory of Depressive Symptomatology (QIDS) change from baseline to endpoint; Better indicated by lower values)												
2 (Poirier 1999, Rush 2006)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	302	293	-	SMD 0.05 higher (0.11 lower to 0.21 higher)	MODERATE	CRITICAL
Remission (ITT) (follow-up 4-14 weeks; assessed with: Number of people scoring <=4/<10 on Hamilton Rating Scale for Depression (HAM-D) or <=5 on Quick Inventory of Depressive Symptomatology (QIDS))												
3 (Lenox-Smith 2008, Poirier 1999, Rush 2006)	randomised trials	serious[1]	very serious[2]	no serious indirectness	serious[3]	reporting bias[4]	145/511 (28.4%)	107/506 (21.1%)	RR 1.48 (0.86 to 2.56)	102 more per 1000 (from 30 fewer to 330 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up 4-14 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D) AND much/very much improved on CGI-I (score 1-2) or at least 50% improvement on Quick Inventory of Depressive Symptomatology (QIDS))												

2 (Poirier 1999, Rush 2006)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	97/311 (31.2%)	81/300 (27%)	RR 1.21 (0.85 to 1.7)	57 more per 1000 (from 40 fewer to 189 more)	LOW	CRITICAL
Discontinuation due to any reason (follow-up 4-12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
2 (Lenox-Smith 2008, Poirier 1999)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[5]	reporting bias[4]	58/261 (22.2%)	50/268 (18.7%)	RR 1.19 (0.85 to 1.67)	35 more per 1000 (from 28 fewer to 125 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up 4-14 weeks; assessed with: Number of participants who dropped out due to adverse events)												
3 (Lenox-Smith 2008, Poirier 1999, Rush 2006)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[6]	none	69/511 (13.5%)	64/506 (12.6%)	RR 1.04 (0.76 to 1.41)	5 more per 1000 (from 30 fewer to 52 more)	LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SNRI: serotonin-norepinephrine reuptake inhibitor
[1] Risk of bias is high or unclear across multiple domains
[2] Considerable heterogeneity
[3] 95% CI crosses thresholds for both clinically important benefit and no effect
[4] Funding from pharmaceutical companies
[5] 95% CI crosses thresholds for both clinically important harm and no effect
[6] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 99: Clinical evidence profile for comparison 30. Switching to SNRI versus switching to bupropion

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Switching to SNRI**	**Switching to bupropion**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology change score (follow-up mean 14 weeks; measured with: Quick Inventory of Depressive Symptomatology (QIDS) change from baseline to endpoint; Better indicated by lower values)												
1 (Rush 2006)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	250	239	-	SMD 0.01 lower (0.19 lower to 0.17 higher)	LOW	CRITICAL
Remission (ITT) (follow-up mean 14 weeks; assessed with: Number of people scoring <=5 on Quick Inventory of Depressive Symptomatology (QIDS))												
1 (Rush 2006)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	62/250 (24.8%)	61/239 (25.5%)	RR 0.97 (0.72 to 1.32)	8 fewer per 1000 (from 71 fewer to 82 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 14 weeks; assessed with: Number of people showing at least 50% improvement on Quick Inventory of Depressive Symptomatology (QIDS))												

1 (Rush 2006)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	70/250 (28%)	62/239 (25.9%)	RR 1.08 (0.81 to 1.45)	21 more per 1000 (from 49 fewer to 117 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 14 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Rush 2006)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	53/250 (21.2%)	65/239 (27.2%)	RR 0.78 (0.57 to 1.07)	60 fewer per 1000 (from 117 fewer to 19 more)	LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SNRI: serotonin-norepinephrine reuptake inhibitor
[1] Risk of bias is high or unclear across multiple domains and abrupt tapering of failed drug
[2] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[3] 95% CI crosses thresholds for both clinically important benefit and no effect

Table 100: Clinical evidence profile for comparison 31. Switching to SNRI versus switching to mirtazapine

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Switching to SNRI**	**Switching to mirtazapine**	**Relative (95% CI)**	**Absolute**		
Remission (ITT) (follow-up mean 8 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Fang 2010)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	21/50 (42%)	20/55 (36.4%)	RR 1.15 (0.72 to 1.86)	55 more per 1000 (from 102 fewer to 313 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 8 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (Fang 2010)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	32/50 (64%)	32/55 (58.2%)	RR 1.1 (0.81 to 1.49)	58 more per 1000 (from 111 fewer to 285 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 8 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Fang 2010)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	9/50 (18%)	10/55 (18.2%)	RR 0.99 (0.44 to 2.24)	2 fewer per 1000 (from 102 fewer to 225 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 8 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Fang 2010)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/50 (0%)	0/55 (0%)	not pooled	not pooled	MODERATE	CRITICAL
Quality of life physical component score (PCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Physical component score; Better indicated by higher values)												
1 (Fang 2010)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	50	55	-	SMD 0.29 higher (0.09 lower to 0.68 higher)	VERY LOW	IMPORTANT

Quality of life mental component score (MCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Mental component score; Better indicated by higher values)												
1 (Fang 2010)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	50	55	-	SMD 0.3 higher (0.08 lower to 0.69 higher)	VERY LOW	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SNRI: serotonin-norepinephrine reuptake inhibitor
[1] Risk of bias is high or unclear across multiple domains
[2] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[3] 95% CI crosses thresholds for both clinically important benefit and no effect

Table 101: Clinical evidence profile for comparison 32. Switching to bupropion versus placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Switching to bupropion**	**Placebo**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology change score (follow-up mean 12 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (GlaxoSmithKline 2009)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[2]	165	157	-	SMD 0.02 higher (0.19 lower to 0.24 higher)	LOW	CRITICAL
Remission (ITT) (follow-up mean 12 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (GlaxoSmithKline 2009)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	reporting bias[2]	40/166 (24.1%)	39/159 (24.5%)	RR 0.98 (0.67 to 1.44)	5 fewer per 1000 (from 81 fewer to 108 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 12 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (GlaxoSmithKline 2009)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	reporting bias[2]	63/166 (38%)	58/159 (36.5%)	RR 1.04 (0.78 to 1.38)	15 more per 1000 (from 80 fewer to 139 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (GlaxoSmithKline 2009)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	reporting bias[2]	67/166 (40.4%)	47/159 (29.6%)	RR 1.37 (1.01 to 1.85)	109 more per 1000 (from 3 more to 251 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 12 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (GlaxoSmithKline 2009)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	reporting bias[2]	39/166 (23.5%)	31/159 (19.5%)	RR 1.21 (0.79 to 1.83)	41 more per 1000 (from 41 fewer to 162 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Rapid tapering of previous treatment
[2] Study run and funded by pharmaceutical company
[3] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[4] 95% CI crosses thresholds for both clinically important harm and no effect

Table 102: Clinical evidence profile for comparison 33. Switching to bupropion versus switching to another antidepressant from same class

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Switching to bupropion**	**Switching to another antidepressant from same class**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology change score (follow-up mean 14 weeks; measured with: Quick Inventory of Depressive Symptomatology (QIDS) change from baseline to endpoint; Better indicated by lower values)												
1 (Rush 2006)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	239	238	-	SMD 0.12 higher (0.06 lower to 0.3 higher)	LOW	CRITICAL
Remission (ITT) (follow-up mean 14 weeks; assessed with: Number of people scoring <=5 on Quick Inventory of Depressive Symptomatology (QIDS))												
1 (Rush 2006)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	61/239 (25.5%)	63/238 (26.5%)	RR 0.96 (0.71 to 1.31)	11 fewer per 1000 (from 77 fewer to 82 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 14 weeks; assessed with: Number of people showing at least 50% improvement on Quick Inventory of Depressive Symptomatology (QIDS))												
1 (Rush 2006)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	62/239 (25.9%)	63/238 (26.5%)	RR 0.98 (0.73 to 1.32)	5 fewer per 1000 (from 71 fewer to 85 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 14 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Rush 2006)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	65/239 (27.2%)	50/238 (21%)	RR 1.29 (0.94 to 1.79)	61 more per 1000 (from 13 fewer to 166 more)	LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains and abrupt tapering of failed drug
[2] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[3] 95% CI crosses thresholds for both clinically important harm and no effect

Table 103: Clinical evidence profile for comparison 34. Augmenting with bupropion versus placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Augmenting with bupropion	Placebo	Relative (95% CI)	Absolute		
Remission (ITT) (follow-up mean 4 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Gulrez 2012)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	18/30 (60%)	7/30 (23.3%)	RR 2.57 (1.26 to 5.24)	366 more per 1000 (from 61 more to 989 more)	MODERATE	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk
[1] Risk of bias is high or unclear across multiple domains

Table 104: Clinical evidence profile for comparison 35. Augmenting with bupropion versus switching to bupropion

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Augmenting with bupropion	Switching to bupropion	Relative (95% CI)	Absolute		
Remission (ITT) (follow-up mean 12 weeks; assessed with: Number of people scoring <=5 on Quick Inventory of Depressive Symptomatology (QIDS))												
1 (Mohamed 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	136/506 (26.9%)	114/511 (22.3%)	RR 1.2 (0.97 to 1.5)	45 more per 1000 (from 7 fewer to 112 more)	MODERATE	CRITICAL
Response (ITT) (follow-up mean 12 weeks; assessed with: Number of people showing at least 50% improvement on Quick Inventory of Depressive Symptomatology (QIDS))												
1 (Mohamed 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	332/506 (65.6%)	319/511 (62.4%)	RR 1.05 (0.96 to 1.15)	31 more per 1000 (from 25 fewer to 94 more)	HIGH	CRITICAL
Discontinuation due to any reason (follow-up mean 12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Mohamed 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	128/506 (25.3%)	158/511 (30.9%)	RR 0.82 (0.67 to 1)	56 fewer per 1000 (from 102 fewer to 0 more)	MODERATE	CRITICAL
Discontinuation due to side effects (follow-up mean 12 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Mohamed 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	37/506 (7.3%)	51/511 (10%)	RR 0.73 (0.49 to 1.1)	27 fewer per 1000 (from 51 fewer to 10 more)	MODERATE	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk

[1] *95% CI crosses thresholds for both clinically important benefit and no effect*

Table 105: Clinical evidence profile for comparison 36. Switching to mirtazapine versus continuing with antidepressant

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Switching to mirtazapine	Continuing with antidepressant	Relative (95% CI)	Absolute		
Depression symptomatology endpoint (follow-up mean 6 weeks; measured with: Patient Health Questionnaire (PHQ-9) or Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
2 (Kato 2018, Xiao 2020)	randomised trials	no serious risk of bias	serious[1]	no serious indirectness	serious[2]	none	618	605	-	SMD 0.21 lower (0.58 lower to 0.17 higher)	LOW	CRITICAL
Depression symptomatology change score (follow-up mean 6 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Xiao 2020)	randomised trials	very serious[3]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[4]	68	68	-	SMD 0.19 lower (0.53 lower to 0.15 higher)	VERY LOW	CRITICAL
Depression symptomatology at 4-month follow-up (follow-up mean 4 months; measured with: Patient Health Questionnaire (PHQ-9); Better indicated by lower values)												
1 (Kato 2018)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	540	538	-	SMD 0.01 higher (0.11 lower to 0.13 higher)	HIGH	CRITICAL
Remission (ITT) (follow-up 6-8 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D) or <=4 on Patient Health Questionnaire (PHQ-9))												
3 (Fang 2010, Kato 2018, Xiao 2020)	randomised trials	no serious risk of bias	serious[1]	no serious indirectness	serious[2]	none	232/681 (34.1%)	185/664 (27.9%)	RR 1.22 (1.04 to 1.43)	61 more per 1000 (from 11 more to 120 more)	LOW	CRITICAL
Remission (ITT) at 4-month follow-up (follow-up mean 4 months; assessed with: Number of people scoring <=4 on Patient Health Questionnaire (PHQ-9))												
1 (Kato 2018)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	262/558 (47%)	245/551 (44.5%)	RR 1.06 (0.93 to 1.2)	27 more per 1000 (from 31 fewer to 89 more)	HIGH	CRITICAL
Response (ITT) (follow-up 6-8 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D) or Patient Health Questionnaire (PHQ-9))												
3 (Fang 2010, Kato 2018, Xiao 2020)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	357/681 (52.4%)	306/664 (46.1%)	RR 1.1 (0.95 to 1.28)	46 more per 1000 (from 23 fewer to 129 more)	MODERATE	CRITICAL
Discontinuation due to any reason (follow-up 6-8 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												

3 (Fang 2010, Kato 2018, Xiao 2020)	randomised trials	serious[5]	no serious inconsistency	no serious indirectness	very serious[6]	reporting bias[7]	30/681 (4.4%)	34/664 (5.1%)	RR 0.85 (0.54 to 1.36)	8 fewer per 1000 (from 24 fewer to 18 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up 6-8 weeks; assessed with: Number of participants who dropped out due to adverse events)												
2 (Fang 2010, Xiao 2020)	randomised trials	serious[5]	no serious inconsistency	no serious indirectness	very serious[6]	reporting bias[7]	3/123 (2.4%)	2/113 (1.8%)	RR 1.19 (0.12 to 11.73)	3 more per 1000 (from 16 fewer to 190 more)	VERY LOW	CRITICAL
Quality of life physical component score (PCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Physical component score; Better indicated by higher values)												
1 (Fang 2010)	randomised trials	very serious[8]	no serious inconsistency	no serious indirectness	serious[9]	none	55	45	-	SMD 0.28 lower (0.67 lower to 0.12 higher)	VERY LOW	IMPORTANT
Quality of life mental component score (MCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Mental component score; Better indicated by higher values)												
1 (Fang 2010)	randomised trials	very serious[8]	no serious inconsistency	no serious indirectness	serious[9]	none	55	45	-	SMD 0.17 lower (0.56 lower to 0.22 higher)	VERY LOW	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Substantial heterogeneity
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] Risk of bias is high across multiple domains
[4] Study partially funded by pharmaceutical company
[5] Statistically significant difference between groups at baseline
[6] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[7] Funding from pharmaceutical companies
[8] Risk of bias is high or unclear across multiple domains
[9] 95% CI crosses thresholds for both clinically important harm and no effect

Table 106: Clinical evidence profile for comparison 37. Augmenting with mirtazapine versus continuing with antidepressant (+/- placebo)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with mirtazapine**	**Continuing with antidepressant (+/- placebo)**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up 4-12 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) or Patient Health Questionnaire (PHQ-9) or Beck Depression Inventory (BDI-II); Better indicated by lower values)												

4 (Carpenter 2002, Kato 2018, Kessler 2018a/2018b, Xiao 2020)	randomised trials	serious[1]	serious[2]	no serious indirectness	no serious imprecision	none	820	837	-	SMD 0.26 lower (0.44 to 0.09 lower)	LOW	CRITICAL
Depression symptomatology change score (follow-up 4-6 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
2 (Carpenter 2002, Xiao 2020)	randomised trials	very serious[1]	very serious[3]	no serious indirectness	serious[4]	reporting bias[5]	79	83	-	SMD 0.52 lower (1.53 lower to 0.48 higher)	VERY LOW	CRITICAL
Depression symptomatology at 4-month follow-up (follow-up mean 4 months; measured with: Patient Health Questionnaire (PHQ-9); Better indicated by lower values)												
1 (Kato 2018)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	520	538	-	SMD 0.07 lower (0.19 lower to 0.05 higher)	HIGH	CRITICAL
Remission (ITT) (follow-up 4-12 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D) or <=4 on Patient Health Questionnaire (PHQ-9) or <10 on Beck Depression Inventory (BDI-II))												
4 (Carpenter 2002, Kato 2018, Kessler 2018a/2018b, Xiao 2020)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	none	290/857 (33.8%)	219/873 (25.1%)	RR 1.3 (1.04 to 1.61)	75 more per 1000 (from 10 more to 153 more)	LOW	CRITICAL
Remission (ITT) at 4-month follow-up (follow-up mean 4 months; assessed with: Number of people scoring <=4 on Patient Health Questionnaire (PHQ-9))												
1 (Kato 2018)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[4]	none	263/537 (49%)	245/551 (44.5%)	RR 1.1 (0.97 to 1.25)	44 more per 1000 (from 13 fewer to 111 more)	MODERATE	CRITICAL
Response (ITT) (follow-up 4-12 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D) or Patient Health Questionnaire (PHQ-9) or Beck Depression Inventory (BDI-II))												
4 (Carpenter 2002, Kato 2018, Kessler 2018a/2018b, Xiao 2020)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	none	422/857 (49.2%)	357/873 (40.9%)	RR 1.19 (1.06 to 1.34)	78 more per 1000 (from 25 more to 139 more)	LOW	CRITICAL
Discontinuation due to any reason (follow-up 4-12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
4 (Carpenter 2002, Kato 2018, Kessler 2018a/2018b, Xiao 2020)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[6]	none	47/857 (5.5%)	50/873 (5.7%)	RR 0.95 (0.65 to 1.4)	3 fewer per 1000 (from 20 fewer to 23 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up 4-6 weeks; assessed with: Number of participants who dropped out due to adverse events)												
2 (Carpenter 2002, Xiao 2020)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[6]	reporting bias[5]	3/79 (3.8%)	2/83 (2.4%)	RR 1.69 (0.29 to 9.93)	17 more per 1000 (from 17 fewer to 215 more)	VERY LOW	CRITICAL

Quality of life endpoint (follow-up mean 12 weeks; measured with: European Quality of Life Questionnaire-5 Dimensions (EQ-5D); Better indicated by higher values)												
1 (Kessler 2018a/2018b)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	213	216	-	SMD 0.04 lower (0.23 lower to 0.15 higher)	LOW	IMPORTANT
Quality of life physical component score (PCS) endpoint (follow-up mean 12 weeks; measured with: 12-item Short-Form Survey (SF-12): Physical component score; Better indicated by higher values)												
1 (Kessler 2018a/2018b)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	208	210	-	SMD 0.14 lower (0.33 lower to 0.05 higher)	LOW	IMPORTANT
Quality of life mental component score (MCS) endpoint (follow-up mean 12 weeks; measured with: 12-item Short-Form Survey (SF-12): Mental component score; Better indicated by higher values)												
1 (Kessler 2018a/2018b)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	208	210	-	SMD 0.29 higher (0.1 to 0.48 higher)	LOW	IMPORTANT
Global functioning endpoint (follow-up mean 4 weeks; measured with: Global Assessment of Function (GAF); Better indicated by higher values)												
1 (Carpenter 2002)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	reporting bias[5]	11	15	-	SMD 0.92 higher (0.1 to 1.75 higher)	VERY LOW	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains
[2] Substantial heterogeneity
[3] Considerable heterogeneity
[4] 95% CI crosses thresholds for both clinically important benefit and no effect
[5] Funding from pharmaceutical companies
[6] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 107: Clinical evidence profile for comparison 38. Augmenting with mirtazapine versus switching to mirtazapine

Quality assessment							**No of patients**		**Effect**		**Quality**	**Importance**
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with mirtazapine**	**Switching to mirtazapine**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 6 weeks; measured with: Patient Health Questionnaire (PHQ-9) or Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												

2 (Kato 2018, Xiao 2020)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	595	618	-	SMD 0.01 lower (0.12 lower to 0.1 higher)	HIGH	CRITICAL
Depression symptomatology change score (follow-up mean 6 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Xiao 2020)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[2]	68	68	-	SMD 0.12 higher (0.22 lower to 0.45 higher)	VERY LOW	CRITICAL
Depression symptomatology at 4-month follow-up (follow-up mean 4 months; measured with: Patient Health Questionnaire (PHQ-9); Better indicated by lower values)												
1 (Kato 2018)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	520	540	-	SMD 0.08 lower (0.2 lower to 0.04 higher)	HIGH	CRITICAL
Remission (ITT) (follow-up mean 6 weeks; assessed with: Number of people scoring <=4 on Patient Health Questionnaire (PHQ-9) or <=7 on Hamilton Rating Scale for Depression (HAM-D))												
2 (Kato 2018, Xiao 2020)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	222/605 (36.7%)	212/626 (33.9%)	RR 1.04 (0.85 to 1.29)	14 more per 1000 (from 51 fewer to 98 more)	MODERATE	CRITICAL
Remission (ITT) at 4-month follow-up (follow-up mean 4 months; assessed with: Number of people scoring <=4 on Patient Health Questionnaire (PHQ-9))												
1 (Kato 2018)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	263/537 (49%)	262/558 (47%)	RR 1.04 (0.92 to 1.18)	19 more per 1000 (from 38 fewer to 85 more)	HIGH	CRITICAL
Response (ITT) (follow-up mean 6 weeks; assessed with: Number of people showing at least 50% improvement on Patient Health Questionnaire (PHQ-9) or Hamilton Rating Scale for Depression (HAM-D))												
2 (Kato 2018, Xiao 2020)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	321/605 (53.1%)	325/626 (51.9%)	RR 1.01 (0.91 to 1.12)	5 more per 1000 (from 47 fewer to 62 more)	HIGH	CRITICAL
Discontinuation due to any reason (follow-up mean 6 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
2 (Kato 2018, Xiao 2020)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	none	19/605 (3.1%)	20/626 (3.2%)	RR 0.95 (0.52 to 1.73)	2 fewer per 1000 (from 15 fewer to 23 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 6 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Xiao 2020)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[2]	2/68 (2.9%)	3/68 (4.4%)	RR 0.67 (0.12 to 3.86)	15 fewer per 1000 (from 39 fewer to 126 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high across multiple domains
[2] Study partially funded by pharmaceutical company
[3] 95% CI crosses threshold for both clinically important benefit and no effect
[4] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 108: Clinical evidence profile for comparison 39. Augmenting with trazodone versus continuing with antidepressant

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Augmenting with trazodone	Continuing with antidepressant	Relative (95% CI)	Absolute		
Remission (ITT) (follow-up mean 8 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Fang 2011)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	20/47 (42.6%)	21/45 (46.7%)	RR 0.91 (0.58 to 1.44)	42 fewer per 1000 (from 196 fewer to 205 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 8 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (Fang 2011)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	29/47 (61.7%)	30/45 (66.7%)	RR 0.93 (0.68 to 1.26)	47 fewer per 1000 (from 213 fewer to 173 more)	VERY LOW	CRITICAL
Quality of life physical component score (PCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Physical component score; Better indicated by higher values)												
1 (Fang 2011)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	47	45	-	SMD 0.26 lower (0.67 lower to 0.15 higher)	LOW	IMPORTANT
Quality of life mental component score (MCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Mental component score; Better indicated by higher values)												
1 (Fang 2011)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	none	47	45	-	SMD 0.2 higher (0.21 lower to 0.61 higher)	LOW	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias was high or unclear across multiple domains
[2] 95% CI crosses thresholds of no effect, and both clinically important benefit and harm
[3] 95% CI crosses thresholds for both clinically important harm and no effect
[4] 95% CI crosses thresholds for both clinically important benefit and no effect

Table 109: Clinical evidence profile for comparison 40. Augmenting with anticonvulsant versus continuing with antidepressant (+/- placebo)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Augmenting with anticonvulsant	Continuing with antidepressant (+/- placebo)	Relative (95% CI)	Absolute		

Depression symptomatology endpoint (follow-up 4-12 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) or Montgomery Asberg Depression Rating Scale (MADRS); Better indicated by lower values)												
8 (Barbee 2011, Li 2009, Li 2015, Mowla 2011, Santos 2008, Wang 2012a, Yang 2016, Zhang 2016)	randomised trials	serious[1]	very serious[2]	no serious indirectness	serious[3]	none	301	298	-	SMD 1.39 lower (2.33 to 0.46 lower)	VERY LOW	CRITICAL
Depression symptomatology change score (follow-up 4-12 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) or Montgomery Asberg Depression Rating Scale (MADRS) change from baseline to endpoint; Better indicated by lower values)												
8 (Barbee 2011, Li 2009, Li 2015, Mowla 2011, Santos 2008, Wang 2012a, Yang 2016, Zhang 2016)	randomised trials	serious[1]	very serious[2]	no serious indirectness	no serious imprecision	none	301	298	-	SMD 1.97 lower (3.07 to 0.87 lower)	VERY LOW	CRITICAL
Remission (ITT) (follow-up mean 8 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Fang 2011)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[4]	none	19/39 (48.7%)	21/45 (46.7%)	RR 1.04 (0.67 to 1.63)	19 more per 1000 (from 154 fewer to 294 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up 4-12 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D) or Montgomery Asberg Depression Rating Scale (MADRS))												
8 (Barbee 2011, Fang 2011, Li 2009, Li 2015, Santos 2008, Wang 2012a, Yang 2016, Zhang 2016)	randomised trials	serious[1]	serious[5]	no serious indirectness	serious[3]	none	149/320 (46.6%)	105/321 (32.7%)	RR 1.44 (0.93 to 2.24)	144 more per 1000 (from 23 fewer to 406 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up 8-10 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
3 (Barbee 2011, Mowla 2011, Santos 2008)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[6]	23/91 (25.3%)	26/92 (28.3%)	RR 0.89 (0.55 to 1.43)	31 fewer per 1000 (from 127 fewer to 122 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up 8-10 weeks; assessed with: Number of participants who dropped out due to adverse events)												
2 (Barbee 2011, Santos 2008)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[6]	9/65 (13.8%)	10/65 (15.4%)	RR 1.12 (0.21 to 5.94)	18 more per 1000 (from 122 fewer to 760 more)	VERY LOW	CRITICAL
Quality of life physical component score (PCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Physical component score; Better indicated by higher values)												

1 (Fang 2011)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[7]	none	39	45	-	SMD 0.21 lower (0.64 lower to 0.22 higher)	LOW	IMPORTANT
Quality of life mental component score (MCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Mental component score; Better indicated by higher values)												
1 (Fang 2011)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	39	45	-	SMD 0.19 higher (0.24 lower to 0.62 higher)	LOW	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains
[2] Considerable heterogeneity
[3] 95% CI crosses thresholds for both clinically important benefit and no effect
[4] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[5] Substantial heterogeneity
[6] Funding from pharmaceutical companies
[7] 95% CI crosses thresholds for both clinically important harm and no effect

Table 110: Clinical evidence profile for comparison 41. Augmenting with anticonvulsant versus lithium

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with anticonvulsant**	**Lithium**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 8 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Schindler 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	17	17	-	SMD 0.31 lower (0.99 lower to 0.36 higher)	LOW	CRITICAL
Depression symptomatology change score (follow-up mean 8 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Schindler 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	17	17	-	SMD 0.81 lower (1.51 to 0.11 lower)	LOW	CRITICAL
Remission (ITT) (follow-up mean 8 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Schindler 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	none	4/17 (23.5%)	3/17 (17.6%)	RR 1.33 (0.35 to 5.08)	58 more per 1000 (from 115 fewer to 720 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 8 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												

1 (Schindler 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	none	9/17 (52.9%)	7/17 (41.2%)	RR 1.29 (0.62 to 2.65)	119 more per 1000 (from 156 fewer to 679 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 8 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Schindler 2007)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	none	2/17 (11.8%)	2/17 (11.8%)	RR 1 (0.16 to 6.3)	0 fewer per 1000 (from 99 fewer to 624 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 8 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Schindler 2007)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/17 (0%)	0/17 (0%)	not pooled	not pooled	HIGH	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 111: Clinical evidence profile for comparison 42. Switching to antipsychotic versus continuing with antidepressant

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Switching to antipsychotic**	**Continuing with antidepressant**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology change score (follow-up 8-12 weeks; measured with: Montgomery Asberg Depression Rating Scale (MADRS) change from baseline to endpoint; Better indicated by lower values)												
3 (Corya 2006, Shelton 2005, Thase 2007)	randomised trials	serious[1]	serious[2]	no serious indirectness	serious[3]	reporting bias[4]	400	329	-	SMD 0.22 higher (0.12 lower to 0.56 higher)	VERY LOW	CRITICAL
Remission (ITT) (follow-up 8-12 weeks; assessed with: Number of people scoring <=8/<=10 on Montgomery Asberg Depression Rating Scale (MADRS))												
3 (Corya 2006, Shelton 2005, Thase 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[4]	56/405 (13.8%)	59/333 (17.7%)	RR 0.79 (0.56 to 1.1)	37 fewer per 1000 (from 78 fewer to 18 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up 8-12 weeks; assessed with: Number of people showing at least 50% improvement on Montgomery Asberg Depression Rating Scale (MADRS))												
3 (Corya 2006, Shelton 2005, Thase 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[4]	94/405 (23.2%)	110/333 (33%)	RR 0.68 (0.48 to 0.96)	106 fewer per 1000 (from 13 fewer to 172 fewer)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up 8-12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												

3 (Corya 2006, Shelton 2005, Thase 2007)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[4]	122/405 (30.1%)	63/333 (18.9%)	RR 1.67 (1.26 to 2.23)	127 more per 1000 (from 49 more to 233 more)	MODERATE	CRITICAL
Discontinuation due to side effects (follow-up 8-12 weeks; assessed with: Number of participants who dropped out due to adverse events)												
3 (Corya 2006, Shelton 2005, Thase 2007)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[4]	51/405 (12.6%)	8/333 (2.4%)	RR 5.34 (2.57 to 11.09)	104 more per 1000 (from 38 more to 242 more)	MODERATE	CRITICAL
Quality of life physical component score (PCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Physical component score; Better indicated by higher values)												
1 (Thase 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[4]	197	203	-	SMD 0.15 lower (0.35 lower to 0.04 higher)	LOW	IMPORTANT
Quality of life mental component score (MCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Mental component score; Better indicated by higher values)												
1 (Thase 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[4]	197	203	-	SMD 0.05 lower (0.25 lower to 0.15 higher)	LOW	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains
[2] Substantial heterogeneity
[3] 95% CI crosses thresholds for both clinically important harm and no effect
[4] Funding from pharmaceutical companies

Table 112: Clinical evidence profile for comparison 43. Switching to combined antipsychotic + SSRI versus continuing with antidepressant

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Switching to combined antipsychotic + SSRI**	**Continuing with antidepressant**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology change score (follow-up 8-12 weeks; measured with: Montgomery Asberg Depression Rating Scale (MADRS) change from baseline to endpoint; Better indicated by lower values)												
2 (Corya 2006, Shelton 2005)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[2]	376	126	-	SMD 0.09 lower (0.3 lower to 0.11 higher)	LOW	CRITICAL

Remission (ITT) (follow-up 8-12 weeks; assessed with: Number of people scoring <=8 on Montgomery Asberg Depression Rating Scale (MADRS))												
2 (Corya 2006, Shelton 2005)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	reporting bias[2]	94/389 (24.2%)	25/127 (19.7%)	RR 1.15 (0.77 to 1.71)	30 more per 1000 (from 45 fewer to 140 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up 8-12 weeks; assessed with: Number of people showing at least 50% improvement on Montgomery Asberg Depression Rating Scale (MADRS))												
2 (Corya 2006, Shelton 2005)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	reporting bias[2]	140/389 (36%)	50/127 (39.4%)	RR 0.85 (0.67 to 1.09)	59 fewer per 1000 (from 130 fewer to 35 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up 8-12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
2 (Corya 2006, Shelton 2005)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	reporting bias[2]	90/389 (23.1%)	23/127 (18.1%)	RR 1.22 (0.69 to 2.16)	40 more per 1000 (from 56 fewer to 210 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up 8-12 weeks; assessed with: Number of participants who dropped out due to adverse events)												
2 (Corya 2006, Shelton 2005)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[4]	reporting bias[2]	39/389 (10%)	3/127 (2.4%)	RR 3.48 (1.06 to 11.44)	59 more per 1000 (from 1 more to 247 more)	LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SSRI: selective serotonin reuptake inhibitor
[1] Risk of bias is high or unclear across multiple domains
[2] Funding from pharmaceutical companies
[3] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[4] 95% CI crosses thresholds for both clinically important harm and no effect

Table 113: Clinical evidence profile for comparison 44. Switching to combined antipsychotic + SSRI versus switch to SSRI-only

Quality assessment							**No of patients**		**Effect**		**Quality**	**Importance**
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Switching to combined antipsychotic + SSRI**	**Switch to SSRI-only**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology change score (follow-up 8-12 weeks; measured with: Montgomery Asberg Depression Rating Scale (MADRS) change from baseline to endpoint; Better indicated by lower values)												
2 (Corya 2006, Shelton 2005)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[2]	376	198	-	SMD 0.12 lower (0.35 lower to 0.1 higher)	LOW	CRITICAL
Remission (ITT) (follow-up 8-12 weeks; assessed with: Number of people scoring <=8 on Montgomery Asberg Depression Rating Scale (MADRS))												

2 (Corya 2006, Shelton 2005)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[2]	94/389 (24.2%)	29/202 (14.4%)	RR 1.46 (0.97 to 2.19)	66 more per 1000 (from 4 fewer to 171 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up 8-12 weeks; assessed with: Number of people showing at least 50% improvement on Montgomery Asberg Depression Rating Scale (MADRS))												
2 (Corya 2006, Shelton 2005)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[2]	140/389 (36%)	60/202 (29.7%)	RR 1.1 (0.81 to 1.5)	30 more per 1000 (from 56 fewer to 149 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up 8-12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
2 (Corya 2006, Shelton 2005)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[2]	90/389 (23.1%)	40/202 (19.8%)	RR 1.12 (0.78 to 1.59)	24 more per 1000 (from 44 fewer to 117 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up 8-12 weeks; assessed with: Number of participants who dropped out due to adverse events)												
2 (Corya 2006, Shelton 2005)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[5]	reporting bias[2]	39/389 (10%)	7/202 (3.5%)	RR 2.41 (1.07 to 5.42)	49 more per 1000 (from 2 more to 153 more)	LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; SSRI: selective serotonin reuptake inhibitor
[1] Risk of bias is high or unclear across multiple domains
[2] Funding from pharmaceutical companies
[3] 95% CI crosses thresholds for both clinically important benefit and no effect
[4] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[5] 95% CI crosses thresholds for both clinically important harm and no effect

Table 114: Clinical evidence profile for comparison 45. Augmenting with antipsychotic versus antidepressant-only or antidepressant + placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with antipsychotic**	**Antidepressant-only or antidepressant + placebo**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up 4-8 weeks; measured with: Montgomery Asberg Depression Rating Scale (MADRS) or Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
5 (Fava 2012/ Mischoulon 2012, Li 2013, Mahmoud 2007, Moica 2018, Song 2007)	randomised trials	serious[1]	very serious[2]	no serious indirectness	serious[3]	none	295	411	-	SMD 0.78 lower (1.24 to	VERY LOW	CRITICAL

										0.32 lower)		
Depression symptomatology change score (follow-up 4-8 weeks; measured with: Montgomery Asberg Depression Rating Scale (MADRS) or Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
20 (Berman 2009, Dunner 2007, Durgam 2016, Earley 2018, Fava 2012/ Mischoulon 2012, Fava 2018, Fava 2019, Hobart 2018a, Hobart 2018b, Kamijima 2013, Kamijima 2018, Li 2013, Moica 2018, Otsuka Pharmaceutical 2015, Otsuka Pharmaceutical 2016, Papakostas 2015, Reeves 2008, Thase 2007, Thase 2015a, Thase 2015b)	randomised trials	serious[1]	serious[4]	no serious indirectness	no serious imprecision	reporting bias[5]	3784	2932	-	SMD 0.33 lower (0.44 to 0.23 lower)	VERY LOW	CRITICAL
Remission (ITT) (follow-up 4-24 weeks; assessed with: Number of people scoring <=10 on Montgomery Asberg Depression Rating Scale (MADRS) or <=7 on Hamilton Rating Scale for Depression (HAM-D))												
28 (Bauer 2009, Bauer 2019, Berman 2007, Berman 2009, Dunner 2007, Durgam 2016, Earley 2018, El-Khalili 2010, Fang 2011, Fava 2012/ Mischoulon 2012, Fava 2018, Fava 2019, Hobart 2018a, Hobart 2018b, Kamijima 2013, Kamijima 2018, Keitner 2009, Li 2013, Lenze 2015, Mahmoud 2007, Marcus 2008, McIntyre 2007, Otsuka Pharmaceutical 2015, Otsuka Pharmaceutical 2016, Papakostas 2015, Thase 2007, Thase 2015a, Thase 2015b)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[5]	1494/5653 (26.4%)	839/4425 (19%)	RR 1.37 (1.23 to 1.52)	70 more per 1000 (from 44 more to 99 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up 4-8 weeks; assessed with: Number of people showing at least 50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) or Hamilton Rating Scale for Depression (HAM-D))												
28 (Bauer 2009, Berman 2007, Berman 2009, Dunner 2007, Durgam 2016, Earley 2018, El-Khalili 2010, Fang 2011, Fava 2012/ Mischoulon 2012, Fava 2018, Fava 2019, Hobart 2018a, Hobart 2018b, Kamijima 2013, Kamijima 2018, Keitner 2009, Li 2013, Mahmoud 2007, Marcus 2008, McIntyre 2007, Otsuka Pharmaceutical 2015,	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[5]	1912/5190 (36.8%)	1025/3964 (25.9%)	RR 1.37 (1.27 to 1.49)	96 more per 1000 (from 70 more to 127 more)	LOW	CRITICAL

Otsuka Pharmaceutical 2016, Papakostas 2015, Reeves 2008, Song 2007, Thase 2007, Thase 2015a, Thase 2015b)												
Discontinuation due to any reason (follow-up 4-24 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
28 (Bauer 2009, Bauer 2019, Berman 2007, Berman 2009, Dunner 2007, Durgam 2016, Earley 2018, El-Khalili 2010, Fava 2012/ Mischoulon 2012, Fava 2018, Fava 2019, Hobart 2018a, Hobart 2018b, Kamijima 2013, Kamijima 2018, Keitner 2009, Lenze 2015, Li 2013, Mahmoud 2007, Marcus 2008, McIntyre 2007, Otsuka Pharmaceutical 2015, Otsuka Pharmaceutical 2016, Papakostas 2015, Reeves 2008, Thase 2007, Thase 2015a, Thase 2015b)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[6]	reporting bias[5]	825/5620 (14.7%)	525/4392 (12%)	RR 1.26 (1.13 to 1.4)	31 more per 1000 (from 16 more to 48 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up 4-24 weeks; assessed with: Number of participants who dropped out due to adverse events)												
27 (Bauer 2009, Bauer 2019, Berman 2007, Berman 2009, Dunner 2007, Durgam 2016, Earley 2018, El-Khalili 2010, Fava 2012/ Mischoulon 2012, Fava 2018, Fava 2019, Hobart 2018a, Hobart 2018b, Kamijima 2013, Kamijima 2018, Keitner 2009, Li 2013, Mahmoud 2007, Marcus 2008, McIntyre 2007, Otsuka Pharmaceutical 2015, Otsuka Pharmaceutical 2016, Papakostas 2015, Reeves 2008, Thase 2007, Thase 2015a, Thase 2015b)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[5]	346/5608 (6.2%)	70/4381 (1.6%)	RR 3.07 (2.36 to 3.99)	33 more per 1000 (from 22 more to 48 more)	MODERATE	CRITICAL
Quality of life endpoint (follow-up mean 6 weeks; measured with: Quality of Life Enjoyment and Satisfaction Questionnaire-short form (Q-LES-Q-SF); Better indicated by higher values)												
1 (Mahmoud 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[5]	101	101	-	SMD 0.47 higher (0.19 to 0.75 higher)	VERY LOW	IMPORTANT
Quality of life change score (follow-up mean 6 weeks; measured with: Quality of Life Enjoyment and Satisfaction Questionnaire-short form (Q-LES-Q-SF) change from baseline to endpoint; Better indicated by higher values)												

2 (Berman 2009, Otsuka Pharmaceutical 2016)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[5]	446	281	-	SMD 0.17 higher (0 to 0.34 higher)	MODERATE	IMPORTANT
Quality of life physical component score (PCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Physical component score; Better indicated by higher values)												
2 (Fang 2011, Thase 2007)	randomised trials	serious[1]	serious[4]	no serious indirectness	no serious imprecision	reporting bias[5]	243	248	-	SMD 0.04 higher (0.33 lower to 0.41 higher)	VERY LOW	IMPORTANT
Quality of life mental component score (MCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Mental component score; Better indicated by higher values)												
2 (Fang 2011, Thase 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[5]	243	248	-	SMD 0.05 higher (0.19 lower to 0.3 higher)	LOW	IMPORTANT
Global functioning change score (follow-up mean 6 weeks; measured with: Social Adaptation Self-evaluation Scale (SASS) change from baseline to endpoint; Better indicated by higher values)												
1 (Kamijima 2018)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[5]	164	149	-	SMD 0.58 higher (0.36 to 0.81 higher)	LOW	IMPORTANT
Functional remission (follow-up mean 24 weeks; assessed with: Number of people scoring <=6 total score on Sheehan Disability Scale (SDS) and all SDS domain scores <=2)												
1 (Bauer 2019)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[7]	reporting bias[5]	68/444 (15.3%)	73/442 (16.5%)	RR 0.93 (0.68 to 1.26)	12 fewer per 1000 (from 53 fewer to 43 more)	VERY LOW	IMPORTANT
Functional impairment endpoint (follow-up mean 6 weeks; measured with: Sheehan Disability Scale (SDS); Better indicated by lower values)												
1 (Mahmoud 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	reporting bias[5]	100	101	-	SMD 0.62 lower (0.9 to 0.34 lower)	VERY LOW	IMPORTANT
Functional impairment change score (follow-up 5-8 weeks; measured with: Sheehan Disability Scale (SDS) change from baseline to endpoint; Better indicated by lower values)												
10 (Berman 2009, Durgam 2016, Fava 2019, Hobart 2018a, Hobart 2018b, Kamijima 2013, Otsuka Pharmaceutical 2015, Otsuka Pharmaceutical	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[5]	2710	1844	-	SMD 0.17 lower (0.24 to 0.11 lower)	LOW	IMPORTANT

2016, Thase 2015a, Thase 2015b)													

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains
[2] Considerable heterogeneity
[3] 95% CI crosses thresholds for both clinically important benefit and no effect
[4] Substantial heterogeneity
[5] Funding from pharmaceutical companies
[6] 95% CI crosses thresholds for both clinically important harm and no effect
[7] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 115: Clinical evidence profile for comparison 46. Augmenting with antipsychotic versus bupropion

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with antipsychotic**	**Bupropion**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology change score (follow-up mean 6 weeks; measured with: Montgomery Asberg Depression Rating Scale (MADRS) change from baseline to endpoint; Better indicated by lower values)												
1 (Cheon 2017)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	56	47	-	SMD 0.48 lower (0.87 to 0.08 lower)	VERY LOW	CRITICAL
Remission (ITT) (follow-up 6-12 weeks; assessed with: Number of people scoring <=10 on Montgomery Asberg Depression Rating Scale (MADRS) or <=5 on Quick Inventory of Depressive Symptomatology (QIDS))												
2 (Cheon 2017, Mohamed 2017)	randomised trials	no serious risk of bias	serious[4]	no serious indirectness	serious[2]	none	177/561 (31.6%)	152/553 (27.5%)	RR 1.25 (0.85 to 1.85)	69 more per 1000 (from 41 fewer to 234 more)	LOW	CRITICAL
Response (ITT) (follow-up 6-12 weeks; assessed with: Number of people showing at least 50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) or Quick Inventory of Depressive Symptomatology (QIDS))												
2 (Cheon 2017, Mohamed 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	409/561 (72.9%)	352/553 (63.7%)	RR 1.17 (1 to 1.38)	108 more per 1000 (from 0 more to 242 more)	MODERATE	CRITICAL
Discontinuation due to any reason (follow-up 6-12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
2 (Cheon 2017, Mohamed 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	113/561 (20.1%)	139/553 (25.1%)	RR 0.8 (0.64 to 1)	50 fewer per 1000 (from 90 fewer to 0 more)	MODERATE	CRITICAL
Discontinuation due to side effects (follow-up 6-12 weeks; assessed with: Number of participants who dropped out due to adverse events)												

2 (Cheon 2017, Mohamed 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	27/561 (4.8%)	37/553 (6.7%)	RR 0.73 (0.45 to 1.18)	18 fewer per 1000 (from 37 fewer to 12 more)	MODERATE	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] Funding from pharmaceutical companies
[4] Substantial heterogeneity

Table 116: Clinical evidence profile for comparison 47. Augmenting with antipsychotic versus lithium

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Augmenting with antipsychotic	Lithium	Relative (95% CI)	Absolute		
Remission (ITT) (follow-up 4-8 weeks; assessed with: Number of people scoring <=8/<=10 on Montgomery Asberg Depression Rating Scale (MADRS) or <=7 on Hamilton Rating Scale for Depression (HAM-D))												
3 (Bauer 2013, Doree 2007, Yoshimura 2014)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	reporting bias[2]	84/261 (32.2%)	65/249 (26.1%)	RR 1.35 (0.82 to 2.22)	91 more per 1000 (from 47 fewer to 318 more)	LOW	CRITICAL
Response (ITT) (follow-up 4-8 weeks; assessed with: Number of people showing at least 50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) or Hamilton Rating Scale for Depression (HAM-D))												
3 (Bauer 2013, Doree 2007, Yoshimura 2014)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	reporting bias[2]	135/261 (51.7%)	111/249 (44.6%)	RR 1.18 (0.98 to 1.41)	80 more per 1000 (from 9 fewer to 183 more)	LOW	CRITICAL
Discontinuation due to any reason (follow-up 4-8 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
3 (Bauer 2013, Doree 2007, Yoshimura 2014)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	reporting bias[2]	36/261 (13.8%)	51/249 (20.5%)	RR 0.71 (0.48 to 1.05)	59 fewer per 1000 (from 107 fewer to 10 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up 4-8 weeks; assessed with: Number of participants who dropped out due to adverse events)												
3 (Bauer 2013, Doree 2007, Yoshimura 2014)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	reporting bias[2]	24/261 (9.2%)	20/249 (8%)	RR 1.16 (0.66 to 2.04)	13 more per 1000 (from 27 fewer to 84 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk
[1] 95% CI crosses thresholds for both clinically important benefit and no effect
[2] Funding from pharmaceutical companies
[3] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 117: Clinical evidence profile for comparison 48. Augmenting with antipsychotic versus switch to antipsychotic

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Augmenting with antipsychotic	Switch to antipsychotic	Relative (95% CI)	Absolute		
Depression symptomatology change score (follow-up mean 8 weeks; measured with: Montgomery Asberg Depression Rating Scale (MADRS) change from baseline to endpoint; Better indicated by lower values)												
1 (Thase 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	198	197	-	SMD 0.38 lower (0.58 to 0.18 lower)	VERY LOW	CRITICAL
Remission (ITT) (follow-up 6-8 weeks; assessed with: Number of people scoring <=10 on Montgomery Asberg Depression Rating Scale (MADRS))												
2 (Bauer 2013, Thase 2007)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	127/431 (29.5%)	82/427 (19.2%)	RR 1.54 (1.14 to 2.07)	104 more per 1000 (from 27 more to 205 more)	LOW	CRITICAL
Response (ITT) (follow-up 6-8 weeks; assessed with: Number of people showing at least 50% improvement on Montgomery Asberg Depression Rating Scale (MADRS))												
2 (Bauer 2013, Thase 2007)	randomised trials	no serious risk of bias	very serious[4]	no serious indirectness	serious[2]	reporting bias[3]	200/431 (46.4%)	165/427 (38.6%)	RR 1.25 (0.84 to 1.88)	97 more per 1000 (from 62 fewer to 340 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up 6-8 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
2 (Bauer 2013, Thase 2007)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	87/431 (20.2%)	121/427 (28.3%)	RR 0.71 (0.56 to 0.9)	82 fewer per 1000 (from 28 fewer to 125 fewer)	LOW	CRITICAL
Discontinuation due to side effects (follow-up 6-8 weeks; assessed with: Number of participants who dropped out due to adverse events)												
2 (Bauer 2013, Thase 2007)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	50/431 (11.6%)	60/427 (14.1%)	RR 0.83 (0.58 to 1.17)	24 fewer per 1000 (from 59 fewer to 24 more)	LOW	CRITICAL
Quality of life physical component score (PCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Physical component score; Better indicated by higher values)												
1 (Thase 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	198	197	-	SMD 0.33 higher (0.13 to 0.53 higher)	VERY LOW	IMPORTANT
Quality of life mental component score (MCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Mental component score; Better indicated by higher values)												
1 (Thase 2007)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[3]	198	197	-	SMD 0.18 higher (0.01 lower to 0.38 higher)	LOW	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference

[1] *Risk of bias is high or unclear across multiple domains*
[2] *95% CI crosses thresholds for both clinically important benefit and no effect*
[3] *Funding from pharmaceutical companies*
[4] *Considerable heterogeneity*

Table 118: Clinical evidence profile for comparison 49. Augmenting with antipsychotic versus switch to bupropion

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Augmenting with antipsychotic	Switch to bupropion	Relative (95% CI)	Absolute		
Remission (ITT) (follow-up mean 12 weeks; assessed with: Number of people scoring <=5 on Quick Inventory of Depressive Symptomatology (QIDS))												
1 (Mohamed 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	146/505 (28.9%)	114/511 (22.3%)	RR 1.3 (1.05 to 1.6)	67 more per 1000 (from 11 more to 134 more)	MODERATE	CRITICAL
Response (ITT) (follow-up mean 12 weeks; assessed with: Number of people showing at least 50% improvement on Quick Inventory of Depressive Symptomatology (QIDS))												
1 (Mohamed 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	375/505 (74.3%)	319/511 (62.4%)	RR 1.19 (1.09 to 1.29)	119 more per 1000 (from 56 more to 181 more)	MODERATE	CRITICAL
Discontinuation due to any reason (follow-up mean 12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Mohamed 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	99/505 (19.6%)	158/511 (30.9%)	RR 0.63 (0.51 to 0.79)	114 fewer per 1000 (from 65 fewer to 152 fewer)	HIGH	CRITICAL
Discontinuation due to side effects (follow-up mean 12 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Mohamed 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	27/505 (5.3%)	51/511 (10%)	RR 0.54 (0.34 to 0.84)	46 fewer per 1000 (from 16 fewer to 66 fewer)	MODERATE	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk
[1] *95% CI crosses thresholds for both clinically important benefit and no effect*

Table 119: Clinical evidence profile for comparison 50. Augmenting with buspirone versus continuing with antidepressant (+/- placebo)

Quality assessment	No of patients	Effect	Quality	Importance

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Augmenting with buspirone	Continuing with antidepressant (+/- placebo)	Relative (95% CI)	Absolute		
Remission (ITT) (follow-up mean 8 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Fang 2011)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	15/46 (32.6%)	21/45 (46.7%)	RR 0.7 (0.42 to 1.18)	140 fewer per 1000 (from 271 fewer to 84 more)	LOW	CRITICAL
Response (ITT) (follow-up 6-8 weeks; assessed with: Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) or showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
2 (Appelberg 2001, Fang 2011)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	43/97 (44.3%)	46/96 (47.9%)	RR 0.9 (0.68 to 1.19)	48 fewer per 1000 (from 153 fewer to 91 more)	LOW	CRITICAL
Quality of life physical component score (PCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Physical component score; Better indicated by higher values)												
1 (Fang 2011)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	46	45	-	SMD 0.06 lower (0.48 lower to 0.35 higher)	MODERATE	IMPORTANT
Quality of life mental component score (MCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Mental component score; Better indicated by higher values)												
1 (Fang 2011)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	46	45	-	SMD 0.08 higher (0.34 lower to 0.49 higher)	MODERATE	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains
[2] 95% CI crosses thresholds for both clinically important harm and no effect

Table 120: Clinical evidence profile for comparison 51. Augmenting with buspirone versus bupropion

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with buspirone**	**Bupropion**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 6 weeks; measured with: Quick Inventory of Depressive Symptomatology (QIDS); Better indicated by lower values)												
1 (Trivedi 2006)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	286	279	-	SMD 0.2 higher (0.04 to 0.37 higher)	MODERATE	CRITICAL
Depression symptomatology change score (follow-up mean 6 weeks; measured with: Quick Inventory of Depressive Symptomatology (QIDS) change from baseline to endpoint; Better indicated by lower values)												

1 (Trivedi 2006)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	286	279	-	SMD 0.17 higher (0.01 to 0.34 higher)	MODERATE	CRITICAL
Remission (ITT) (follow-up mean 6 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Trivedi 2006)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	86/286 (30.1%)	83/279 (29.7%)	RR 1.01 (0.79 to 1.3)	3 more per 1000 (from 62 fewer to 89 more)	LOW	CRITICAL
Response (ITT) (follow-up mean 6 weeks; assessed with: Number of people showing at least 50% improvement on Quick Inventory of Depressive Symptomatology (QIDS))												
1 (Trivedi 2006)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	77/286 (26.9%)	88/279 (31.5%)	RR 0.85 (0.66 to 1.1)	47 fewer per 1000 (from 107 fewer to 32 more)	MODERATE	CRITICAL
Discontinuation due to side effects (follow-up mean 6 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Trivedi 2006)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[3]	none	59/286 (20.6%)	35/279 (12.5%)	RR 1.64 (1.12 to 2.41)	80 more per 1000 (from 15 more to 177 more)	MODERATE	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains
[2] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[3] 95% CI crosses thresholds for both clinically important harm and no effect

Table 121: Clinical evidence profile for comparison 52. Augmenting with methylphenidate versus placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with methylphenidate**	**Placebo**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology change score (follow-up mean 5 weeks; measured with: Montgomery Asberg Depression Rating Scale (MADRS) change from baseline to endpoint; Better indicated by lower values)												
1 (Ravindran 2008a)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[2]	72	72	-	SMD 0.06 higher (0.27 lower to 0.38 higher)	VERY LOW	CRITICAL
Remission (ITT) (follow-up mean 4 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (Patkar 2006)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	reporting bias[2]	4/30 (13.3%)	1/30 (3.3%)	RR 4 (0.47 to 33.73)	100 more per 1000 (from 18 fewer to 1000 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up 4-5 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D) or Montgomery Asberg Depression Rating Scale (MADRS))												
2 (Patkar 2006, Ravindran 2008a)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	reporting bias[2]	46/103 (44.7%)	37/102 (36.3%)	RR 1.21 (0.87 to 1.68)	76 more per 1000 (from 47 fewer to 247 more)	VERY LOW	CRITICAL

Discontinuation due to any reason (follow-up mean 5 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Ravindran 2008a)	randomised trials	serious[5]	no serious inconsistency	no serious indirectness	serious[6]	reporting bias[2]	11/73 (15.1%)	4/72 (5.6%)	RR 2.71 (0.91 to 8.12)	95 more per 1000 (from 5 fewer to 396 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up 4-5 weeks; assessed with: Number of participants who dropped out due to adverse events)												
2 (Patkar 2006, Ravindran 2008a)	randomised trials	no serious risk of bias	serious[7]	no serious indirectness	very serious[3]	reporting bias[2]	8/103 (7.8%)	2/102 (2%)	RR 2.92 (0.21 to 40.65)	38 more per 1000 (from 15 fewer to 777 more)	VERY LOW	CRITICAL

Abbreviations: CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains
[2] Funding from pharmaceutical companies
[3] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[4] 95% CI crosses thresholds for both clinically important benefit and no effect
[5] Statistically significant group difference at baseline
[6] 95% CI crosses thresholds for both clinically important harm and no effect
[7] Substantial heterogeneity

Table 122: Clinical evidence profile for comparison 53. Augmenting with lithium versus continuing with antidepressant (+/- placebo)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with lithium**	**Continuing with antidepressant (+/- placebo)**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up 2-3 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) or Montgomery Asberg Depression Rating Scale (MADRS); Better indicated by lower values)												
2 (Joffe 1993, Stein 1993)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	33	34	-	SMD 0.23 lower (0.71 lower to 0.25 higher)	LOW	CRITICAL
Depression symptomatology change score (follow-up 2-52 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) or Montgomery Asberg Depression Rating Scale (MADRS) or Quick Inventory of Depressive Symptomatology (QIDS) change from baseline to endpoint; Better indicated by lower values)												
3 (Girlanda 2014, Joffe 1993, Stein 1993)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	60	56	-	SMD 0.26 lower (0.76 lower to 0.23 higher)	LOW	CRITICAL
Remission (ITT) (follow-up mean 3 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D) AND responding (at least 50% improvement on HAM-D))												
1 (Joffe 1993)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	none	6/18 (33.3%)	2/16 (12.5%)	RR 2.67 (0.62 to 11.39)	209 more per 1000 (from 47	LOW	CRITICAL

										fewer to 1000 more)		
Response (ITT) (follow-up 1-6 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
2 (Baumann 1996, Nierenberg 2003a)	randomised trials	serious[1]	serious[4]	no serious indirectness	very serious[3]	reporting bias[5]	8/28 (28.6%)	5/31 (16.1%)	RR 1.72 (0.27 to 11.05)	116 more per 1000 (from 118 fewer to 1000 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up 2-52 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
4 (Girlanda 2014, Joffe 1993, Nierenberg 2003a, Stein 1993)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	none	5/81 (6.2%)	7/78 (9%)	RR 0.67 (0.22 to 2.03)	30 fewer per 1000 (from 70 fewer to 92 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up 2-3 weeks; assessed with: Number of participants who dropped out due to adverse events)												
2 (Joffe 1993, Stein 1993)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	none	1/34 (2.9%)	0/34 (0%)	RR 2.68 (0.12 to 61.58)	-	LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias was high or unclear across multiple domains
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[4] Substantial heterogeneity
[5] Funding from pharmaceutical companies

Table 123: Clinical evidence profile for comparison 54. Augmenting with lithium versus switch to antipsychotic

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with lithium**	**Switch to antipsychotic**	**Relative (95% CI)**	**Absolute**		
Remission (ITT) (follow-up mean 6 weeks; assessed with: Number of people scoring <=10 on Montgomery Asberg Depression Rating Scale (MADRS))												
1 (Bauer 2013)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	60/229 (26.2%)	53/228 (23.2%)	RR 1.13 (0.82 to 1.55)	30 more per 1000 (from 42 fewer to 128 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 6 weeks; assessed with: Number of people showing at least 50% improvement on Montgomery Asberg Depression Rating Scale (MADRS))												
1 (Bauer 2013)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	reporting bias[3]	102/229 (44.5%)	114/228 (50%)	RR 0.89 (0.73 to 1.08)	55 fewer per 1000 (from 135 fewer to 40 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 6 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												

1 (Bauer 2013)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[5]	reporting bias[3]	47/229 (20.5%)	49/228 (21.5%)	RR 0.95 (0.67 to 1.36)	11 fewer per 1000 (from 71 fewer to 77 more)	VERY LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 6 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Bauer 2013)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	reporting bias[3]	18/229 (7.9%)	28/228 (12.3%)	RR 0.64 (0.36 to 1.12)	44 fewer per 1000 (from 79 fewer to 15 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk
[1] Rapid switch from failed drug for quetiapine monotherapy arm
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] Study funded by pharmaceutical company
[4] 95% CI crosses thresholds for both clinically important harm and no effect
[5] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 124: Clinical evidence profile for comparison 55. Augmenting with lithium versus augmenting with a psychological intervention

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with lithium**	**Augmenting with a psychological intervention**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 8 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Kennedy 2003)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	19	20	-	SMD 0.41 lower (1.05 lower to 0.22 higher)	MODERATE	CRITICAL
Depression symptomatology change score (follow-up mean 8 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Kennedy 2003)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	19	20	-	SMD 0.42 lower (1.06 lower to 0.21 higher)	MODERATE	CRITICAL
Depression symptomatology at 1-month follow-up (follow-up mean 1 months; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Kennedy 2003)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	19	20	-	SMD 0.65 lower (1.29 lower to 0 higher)	MODERATE	CRITICAL
Remission (ITT) (follow-up mean 8 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												

1 (Kennedy 2003)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	8/21 (38.1%)	6/23 (26.1%)	RR 1.46 (0.61 to 3.51)	120 more per 1000 (from 102 fewer to 655 more)	LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 8 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Kennedy 2003)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	3/21 (14.3%)	3/23 (13%)	RR 1.1 (0.25 to 4.84)	13 more per 1000 (from 98 fewer to 501 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up mean 8 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Kennedy 2003)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	1/21 (4.8%)	0/23 (0%)	RR 3.27 (0.14 to 76.21)	-	LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] 95% CI crosses thresholds for both clinically important benefit and no effect
[2] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 125: Clinical evidence profile for comparison 56. Augmenting with lithium versus augmenting with TCA

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with lithium**	**Augmenting with TCA**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
2 (Fava 1994a, Fava 2002)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	48	46	-	SMD 0.32 lower (0.73 lower to 0.09 higher)	LOW	CRITICAL
Depression symptomatology change score (follow-up mean 4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
2 (Fava 1994a, Fava 2002)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	46	48	-	SMD 0.1 higher (0.31 lower to 0.51 higher)	LOW	CRITICAL
Remission (ITT) (follow-up mean 4 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
2 (Fava 1994a, Fava 2002)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[4]	none	12/48 (25%)	13/46 (28.3%)	RR 0.88 (0.45 to 1.74)	34 fewer per 1000 (from 155 fewer to 209 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 4 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
2 (Fava 1994a, Fava 2002)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	none	7/48 (14.6%)	8/46 (17.4%)	RR 0.83 (0.33 to 2.11)	30 fewer per 1000 (from 117 fewer to 193 more)	LOW	CRITICAL

Discontinuation due to side effects (follow-up mean 4 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Fava 1994a)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	reporting bias[5]	1/14 (7.1%)	2/12 (16.7%)	RR 0.43 (0.04 to 4.16)	95 fewer per 1000 (from 160 fewer to 527 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; TCA: tricyclic antidepressant
[1] Risk of bias high or unclear across multiple domains
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] 95% CI crosses thresholds for both clinically important harm and no effect
[4] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[5] Study partially funded by pharmaceutical company

Table 126: Clinical evidence profile for comparison 57. Augmenting with omega-3 fatty acids versus placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with omega-3 fatty acids**	**Placebo**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up 4-12 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
3 (Jahanggard 2018, Mozaffari-Khosravi 2013, Nemets 2002)	randomised trials	serious[1]	very serious[2]	no serious indirectness	serious[3]	none	76	56	-	SMD 1.73 lower (3.59 lower to 0.12 higher)	VERY LOW	CRITICAL
Depression symptomatology change score (follow-up 4-12 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
3 (Jahanggard 2018, Mozaffari-Khosravi 2013, Nemets 2002)	randomised trials	serious[1]	very serious[2]	no serious indirectness	serious[3]	none	76	56	-	SMD 1.65 lower (3.02 to 0.27 lower)	VERY LOW	CRITICAL
Remission (ITT) (follow-up mean 12 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Mozaffari-Khosravi 2013)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[4]	none	5/54 (9.3%)	0/27 (0%)	RR 5.6 (0.32 to 97.69)	-	VERY LOW	CRITICAL
Response (ITT) (follow-up 4-12 weeks; assessed with: Number of people showing at least 50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) or at least 30% or 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
3 (Mozaffari-Khosravi 2013, Nemets 2002, Peet 2002)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[4]	none	28/116 (24.1%)	5/54 (9.3%)	RR 2.49 (0.77 to 8.06)	138 more per 1000 (from 21 fewer to 654 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up 4-12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												

4 (Jahangard 2018, Mozaffari-Khosravi 2013, Nemets 2002, Peet 2002)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	none	19/141 (13.5%)	11/80 (13.8%)	RR 0.8 (0.41 to 1.56)	27 fewer per 1000 (from 81 fewer to 77 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up 4-12 weeks; assessed with: Number of participants who dropped out due to adverse events)												
4 (Jahangard 2018, Mozaffari-Khosravi 2013, Nemets 2002, Peet 2002)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	none	6/141 (4.3%)	5/80 (6.3%)	RR 0.57 (0.18 to 1.73)	27 fewer per 1000 (from 51 fewer to 46 more)	LOW	CRITICAL
Sleeping difficulties endpoint (follow-up mean 12 weeks; measured with: Insomnia Severity Index (ISI); Better indicated by lower values)												
1 (Jahangard 2018)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	25	25	-	SMD 3.36 lower (4.24 to 2.47 lower)	HIGH	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains
[2] Considerable heterogeneity
[3] 95% CI crosses thresholds for both clinically important benefit and no effect
[4] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 127: Clinical evidence profile for comparison 58. Augmenting with thyroid hormone versus continuing with antidepressant (+/- placebo)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with thyroid hormone**	**Continuing with antidepressant (+/- placebo)**	**Relative (95% CI)**	**Absolute**		
Depression symptoms endpoint (follow-up mean 2 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Joffe 1993)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	17	16	-	SMD 0.53 lower (1.22 lower to 0.17 higher)	MODERATE	CRITICAL
Depression symptoms change score (follow-up mean 2 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Joffe 1993)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	17	16	-	SMD 0.78 lower (1.5 to 0.07 lower)	MODERATE	CRITICAL
Remission (ITT) (follow-up 2-8 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												

2 (Fang 2011, Joffe 1993)	randomised trials	serious[2]	serious[3]	no serious indirectness	very serious[4]	none	25/65 (38.5%)	23/61 (37.7%)	RR 1.39 (0.35 to 5.53)	147 more per 1000 (from 245 fewer to 1000 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 8 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (Fang 2011)	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[5]	none	28/48 (58.3%)	30/45 (66.7%)	RR 0.88 (0.64 to 1.2)	80 fewer per 1000 (from 240 fewer to 133 more)	LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 2 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Joffe 1993)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/17 (0%)	0/16 (0%)	not pooled	not pooled	HIGH	CRITICAL
Discontinuation due to side effects (follow-up mean 2 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Joffe 1993)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/17 (0%)	0/16 (0%)	not pooled	not pooled	HIGH	CRITICAL
Quality of life physical component score (PCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Physical component score; Better indicated by higher values)												
1 (Fang 2011)	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[5]	none	48	45	-	SMD 0.12 lower (0.53 lower to 0.28 higher)	LOW	IMPORTANT
Quality of life mental component score (MCS) change score (follow-up mean 8 weeks; measured with: 36-item Short-Form Survey (SF-36): Mental component score; Better indicated by higher values)												
1 (Fang 2011)	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	48	45	-	SMD 0.02 lower (0.42 lower to 0.39 higher)	MODERATE	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] 95% CI crosses thresholds for both clinically important benefit and no effect
[2] Risk of bias is high or unclear across multiple domains
[3] Substantial heterogeneity
[4] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[5] 95% CI crosses thresholds for both clinically important harm and no effect

Table 128: Clinical evidence profile for comparison 59. Augmenting with thyroid hormone versus augmenting with lithium

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Augmenting with thyroid hormone	Augmenting with lithium	Relative (95% CI)	Absolute		

Depression symptomatology endpoint (follow-up 2-14 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) or Quick Inventory of Depressive Symptomatology (QIDS); Better indicated by lower values)												
2 (Joffe 1993, Nierenberg 2006)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	90	86	-	SMD 0.33 lower (0.63 to 0.03 lower)	VERY LOW	CRITICAL
Depression symptomatology change score (follow-up 2-14 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) or Quick Inventory of Depressive Symptomatology (QIDS) change from baseline to endpoint; Better indicated by lower values)												
2 (Joffe 1993, Nierenberg 2006)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	90	86	-	SMD 0.15 lower (0.45 lower to 0.14 higher)	LOW	CRITICAL
Remission (ITT) (follow-up 2-14 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D) or <=5 on Quick Inventory of Depressive Symptomatology (QIDS))												
2 (Joffe 1993, Nierenberg 2006)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	25/90 (27.8%)	15/87 (17.2%)	RR 1.58 (0.91 to 2.77)	100 more per 1000 (from 16 fewer to 305 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up mean 14 weeks; assessed with: Number of people showing at least 50% improvement on Quick Inventory of Depressive Symptomatology (QIDS))												
1 (Nierenberg 2006)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	none	17/73 (23.3%)	11/69 (15.9%)	RR 1.46 (0.74 to 2.89)	73 more per 1000 (from 41 fewer to 301 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up mean 2 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Joffe 1993)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	none	0/17 (0%)	1/18 (5.6%)	RR 0.35 (0.02 to 8.09)	36 fewer per 1000 (from 54 fewer to 394 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up 2-14 weeks; assessed with: Number of participants who dropped out due to adverse events)												
2 (Joffe 1993, Nierenberg 2006)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	7/90 (7.8%)	17/87 (19.5%)	RR 0.41 (0.18 to 0.91)	115 fewer per 1000 (from 18 fewer to 160 fewer)	LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 129: Clinical evidence profile for comparison 60. Switching to ECT versus switching to paroxetine

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Switching to ECT	Switching to paroxetine	Relative (95% CI)	Absolute		

Depression symptomatology endpoint (follow-up 2-4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Folkerts 1997)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	21	18	-	SMD 1.35 lower (2.06 to 0.65 lower)	LOW	CRITICAL
Depression symptomatology change score (follow-up 2-4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Folkerts 1997)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	21	18	-	SMD 1.61 lower (2.34 to 0.87 lower)	LOW	CRITICAL
Response (ITT) (follow-up 2-4 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
1 (Folkerts 1997)	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	15/21 (71.4%)	5/19 (26.3%)	RR 2.71 (1.22 to 6.04)	450 more per 1000 (from 58 more to 1000 more)	VERY LOW	CRITICAL
Discontinuation due to any reason (follow-up 2-4 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
1 (Folkerts 1997)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	none	0/21 (0%)	1/19 (5.3%)	RR 0.3 (0.01 to 7.02)	37 fewer per 1000 (from 52 fewer to 317 more)	LOW	CRITICAL
Discontinuation due to side effects (follow-up 2-4 weeks; assessed with: Number of participants who dropped out due to adverse events)												
1 (Folkerts 1997)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/21 (0%)	0/19 (0%)	not pooled	not pooled	HIGH	CRITICAL

CI: confidence interval; ECT: electroconvulsive therapy; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains and rapid tapering of prior antidepressant treatment
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 130: Clinical evidence profile for comparison 61. Augmenting with ECT versus continuing with antidepressant

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with ECT**	**Continuing with antidepressant**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Haghighi 2013)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	20	20	-	SMD 0.08 higher (0.54 lower to 0.7 higher)	VERY LOW	CRITICAL
Depression symptomatology change score (follow-up mean 4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Haghighi 2013)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	20	20	-	SMD 0.6 lower (1.23 lower to 0.04 higher)	LOW	CRITICAL

CI: confidence interval; ECT: electroconvulsive therapy; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains
[2] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[3] 95% CI crosses thresholds for both clinically important benefit and no effect

Table 131: Clinical evidence profile for comparison 62. Augmenting with ECT versus augmenting with exercise

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with ECT**	**Augmenting with exercise**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Salehi 2016)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	none	20	20	-	SMD 0.12 higher (0.5 lower to 0.74 higher)	LOW	CRITICAL
Depression symptomatology change score (follow-up mean 4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Salehi 2016)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	20	20	-	SMD 0.18 lower (0.81 lower to 0.44 higher)	MODERATE	CRITICAL
Remission (ITT) (follow-up mean 4 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Salehi 2016)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	none	2/20 (10%)	2/20 (10%)	RR 1 (0.16 to 6.42)	0 fewer per 1000 (from 84 fewer to 542 more)	LOW	CRITICAL

CI: confidence interval; ECT: electroconvulsive therapy; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[2] 95% CI crosses thresholds for both clinically important benefit and no effect

Table 132: Clinical evidence profile for comparison 63. Augmenting with ECT + exercise versus augmenting with exercise

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with ECT + exercise**	**Augmenting with exercise**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												

1 (Salehi 2016)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	20	20	-	SMD 0.99 lower (1.65 to 0.33 lower)	MODERATE	CRITICAL
Depression symptomatology change score (follow-up mean 4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Salehi 2016)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	20	20	-	SMD 1.84 lower (2.59 to 1.09 lower)	HIGH	CRITICAL
Remission (ITT) (follow-up mean 4 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Salehi 2016)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	13/20 (65%)	2/20 (10%)	RR 6.5 (1.68 to 25.16)	550 more per 1000 (from 68 more to 1000 more)	HIGH	CRITICAL

CI: confidence interval; ECT: electroconvulsive therapy; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] 95% CI crosses thresholds for both clinically important benefit and no effect

Table 133: Clinical evidence profile for comparison 64. Augmenting with exercise versus TAU

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with exercise**	**TAU**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 3 weeks; measured with: Montgomery Asberg Depression Rating Scale (MADRS); Better indicated by lower values)												
1 (Ho 2014)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	26	26	-	SMD 0.59 lower (1.15 to 0.04 lower)	MODERATE	CRITICAL
Depression symptomatology change score (follow-up 3-10 weeks; measured with: Montgomery Asberg Depression Rating Scale (MADRS) change from baseline to endpoint; Better indicated by lower values)												
2 (Danielsson 2014, Ho 2014)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	48	46	-	SMD 0.68 lower (1.1 to 0.26 lower)	MODERATE	CRITICAL
Remission (ITT) (follow-up 3-10 weeks; assessed with: Number of people scoring <=10 on Montgomery Asberg Depression Rating Scale (MADRS))												
2 (Danielsson 2014, Ho 2014)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	21/48 (43.8%)	10/46 (21.7%)	RR 2.03 (1.09 to 3.79)	224 more per 1000 (from 20 more to 607 more)	MODERATE	CRITICAL
Response (ITT) (follow-up mean 10 weeks; assessed with: Number of people showing at least 50% improvement on Montgomery Asberg Depression Rating Scale (MADRS))												
1 (Danielsson 2014)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	9/22 (40.9%)	5/20 (25%)	RR 1.64 (0.66 to 4.07)	160 more per 1000 (from 85 fewer to 768 more)	LOW	CRITICAL

Discontinuation due to any reason (follow-up 3-10 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
2 (Danielsson 2014, Ho 2014)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	11/48 (22.9%)	9/46 (19.6%)	RR 1.18 (0.54 to 2.59)	35 more per 1000 (from 90 fewer to 311 more)	LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference; TAU: treatment as usual
[1] 95% CI crosses thresholds for both clinically important benefit and no effect
[2] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 134: Clinical evidence profile for comparison 65. Augmenting with exercise versus attention-placebo

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with exercise**	**Attention-placebo**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 10 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Lavretsky 2011)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	33	35	-	SMD 0.4 lower (0.88 lower to 0.08 higher)	MODERATE	CRITICAL
Depression symptomatology change score (follow-up mean 12 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Mota-Pereira 2011)	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[3]	19	10	-	SMD 5.47 lower (7.17 to 3.77 lower)	LOW	CRITICAL
Remission (ITT) (follow-up 10-12 weeks; assessed with: Number of people scoring <=7 or <7 on Hamilton Rating Scale for Depression (HAM-D))												
2 (Lavretsky 2011, Mota-Pereira 2011)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	none	26/58 (44.8%)	18/48 (37.5%)	RR 1.5 (0.47 to 4.77)	188 more per 1000 (from 199 fewer to 1000 more)	LOW	CRITICAL
Response (ITT) (follow-up 10-12 weeks; assessed with: Number of people showing at least 30% or 50% improvement on Hamilton Rating Scale for Depression (HAM-D))												
2 (Mather 2002, Mota-Pereira 2011)	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[1]	none	27/65 (41.5%)	14/54 (25.9%)	RR 1.7 (1.03 to 2.81)	181 more per 1000 (from 8 more to 469 more)	LOW	CRITICAL
Discontinuation due to any reason (follow-up 10-12 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
3 (Lavretsky 2011, Mather 2002, Mota-Pereira 2011)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[4]	none	6/101 (5.9%)	3/91 (3.3%)	RR 1.53 (0.4 to 5.86)	17 more per 1000 (from 20 fewer to 160 more)	LOW	CRITICAL
Global functioning change score (follow-up mean 12 weeks; measured with: Global Assessment of Function (GAF); Better indicated by higher values)												

1 (Mota-Pereira 2011)	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[3]	19	10	-	SMD 6.15 higher (4.28 to 8.02 higher)	LOW	IMPORTANT
Sleeping difficulties endpoint (follow-up mean 10 weeks; measured with: Pittsburgh Sleep Quality Index (PSQI); Better indicated by lower values)												
1 (Lavretsky 2011)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	33	35	-	SMD 0.25 lower (0.72 lower to 0.23 higher)	MODERATE	IMPORTANT

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] 95% CI crosses thresholds for both clinically important benefit and no effect
[2] Risk of bias is high or unclear across multiple domains
[3] Study partially funded by pharmaceutical company
[4] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm

Table 135: Clinical evidence profile for comparison 66. Augmenting with exercise + ECT versus augmenting with ECT

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Augmenting with exercise + ECT**	**Augmenting with ECT**	**Relative (95% CI)**	**Absolute**		
Depression symptomatology endpoint (follow-up mean 4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D); Better indicated by lower values)												
1 (Salehi 2016)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	20	20	-	SMD 1.13 lower (1.81 to 0.46 lower)	MODERATE	CRITICAL
Depression symptomatology change score (follow-up mean 4 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Salehi 2016)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	20	20	-	SMD 1.45 lower (2.15 to 0.74 lower)	HIGH	CRITICAL
Remission (ITT) (follow-up mean 4 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D))												
1 (Salehi 2016)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	13/20 (65%)	2/20 (10%)	RR 6.5 (1.68 to 25.16)	550 more per 1000 (from 68 more to 1000 more)	HIGH	CRITICAL

CI: confidence interval; ECT: electroconvulsive therapy; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] 95% CI crosses thresholds for both clinically important benefit and no effect

Table 136: Clinical evidence profile for comparison 67. Augmenting with yoga versus continuing with antidepressant (+/- waitlist or attention-placebo)

Quality assessment	No of patients	Effect	Quality	Importance

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Augmenting with yoga	Continuing with antidepressant (+/- waitlist or attention-placebo)	Relative (95% CI)	Absolute		
Depression symptomatology change score (follow-up mean 8 weeks; measured with: Hamilton Rating Scale for Depression (HAM-D) change from baseline to endpoint; Better indicated by lower values)												
1 (Sharma 2017)	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	13	12	-	SMD 1.49 lower (2.39 to 0.58 lower)	HIGH	CRITICAL
Remission (ITT) (follow-up 8-10 weeks; assessed with: Number of people scoring <=7 on Hamilton Rating Scale for Depression (HAM-D) or <=5 on Quick Inventory of Depressive Symptomatology (QIDS))												
2 (Sharma 2017, Uebelacker 2017)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	21/76 (27.6%)	12/71 (16.9%)	RR 1.58 (0.84 to 3)	98 more per 1000 (from 27 fewer to 338 more)	LOW	CRITICAL
Remission (ITT) at 3-month follow-up (follow-up mean 3 months; assessed with: Number of people scoring <=5 on Quick Inventory of Depressive Symptomatology (QIDS))												
1 (Uebelacker 2017)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	19/63 (30.2%)	11/59 (18.6%)	RR 1.62 (0.84 to 3.11)	116 more per 1000 (from 30 fewer to 393 more)	LOW	CRITICAL
Remission (ITT) at 6-month follow-up (follow-up mean 6 months; assessed with: Number of people scoring <=5 on Quick Inventory of Depressive Symptomatology (QIDS))												
1 (Uebelacker 2017)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	none	19/63 (30.2%)	14/59 (23.7%)	RR 1.27 (0.7 to 2.3)	64 more per 1000 (from 71 fewer to 308 more)	VERY LOW	CRITICAL
Response (ITT) (follow-up 8-10 weeks; assessed with: Number of people showing at least 50% improvement on Hamilton Rating Scale for Depression (HAM-D) or Quick Inventory of Depressive Symptomatology (QIDS))												
2 (Sharma 2017, Uebelacker 2017)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	none	27/76 (35.5%)	14/71 (19.7%)	RR 2.06 (0.68 to 6.19)	209 more per 1000 (from 63 fewer to 1000 more)	VERY LOW	CRITICAL
Response (ITT) at 3-month follow-up (follow-up mean 3 months; assessed with: Number of people showing at least 50% improvement on Quick Inventory of Depressive Symptomatology (QIDS))												
1 (Uebelacker 2017)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	22/63 (34.9%)	13/59 (22%)	RR 1.58 (0.88 to 2.85)	128 more per 1000 (from 26 fewer to 408 more)	LOW	CRITICAL
Response (ITT) at 6-month follow-up (follow-up mean 6 months; assessed with: Number of people showing at least 50% improvement on Quick Inventory of Depressive Symptomatology (QIDS))												
1 (Uebelacker 2017)	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	23/63 (36.5%)	14/59 (23.7%)	RR 1.54 (0.88 to 2.7)	128 more per 1000 (from 28	LOW	CRITICAL

										fewer to 403 more)		
Discontinuation due to any reason (follow-up 8-10 weeks; assessed with: Number of participants who dropped out for any reason (including adverse events))												
2 (Sharma 2017, Uebelacker 2017)	randomised trials	no serious risk of bias	serious[4]	no serious indirectness	very serious[3]	none	7/76 (9.2%)	13/71 (18.3%)	RR 0.88 (0.08 to 9.88)	22 fewer per 1000 (from 168 fewer to 1000 more)	VERY LOW	CRITICAL

CI: confidence interval; ITT: intention to treat; RR: relative risk; SMD: standardised mean difference
[1] Risk of bias is high or unclear across multiple domains
[2] 95% CI crosses thresholds for both clinically important benefit and no effect
[3] 95% CI crosses thresholds for no effect, and both clinically important benefit and harm
[4] Substantial heterogeneity

Click here to enter text.