Table 1. Summary of results of clinical effectiveness and safety

PRIMARY OUTCOMES										
Relapse (9 studies, 3 continuation, 6 maintenance; 8 of 9 - high risk of bias)										
	ADM vs Placebo (95% CI) [study n]	**CBT vs Placebo (95% CI) [study n]**	**ADM vs CBT (95% CI) [study n]**	**ADM vs. ADM + CBT (95% CI) [study n]**	**CBT vs CBT + ADM (95% CI) [study n]**	**ADM vs TAU (95% CI) [study n]**	**CBT vs TAU (95% CI) [study n]**	**ADM vs WL (95% CI) [study n]**	**CBT vs. WL (95% CI) [study n]**	**Comments**
Continuation (13 – 24 weeks)	**RR 0.49*** (0.34 to 0.68) AGM, 24 wks **RR 0.49** (0.29 to 0.83) AGM, FXT, 18 wks [2]	no studies	RR 2.48 (0.69 to 8.87) (trend favours CBT) [1]	no studies	no studies	no studies	no studies	no studies	no studies	*similar for severe MDD subgroup
	Significant	-	no significant difference	-	-	-	-	-	-	
Maintenance (≥ 25 weeks)	RR 0.84 (0.44 to 1.59)* [2] *Post-hoc* 21% vs 31% relapse in more severe MDD	16% relapse CBT vs 25% relapse placebo (44 weeks) [1]	RR 1.19 95% CI, 0.50 to 2.80 (96 wks) RR 0.86 95% CI, 0.31 to 2.40 (44 wks) [2]	80 relapses in 49 patients (ADM) vs. 66 relapses in 44 patients (ADM + CBT) [1]	no studies	no studies	no studies	no studies	no studies	*studies inconsistent
	Significant	no significant difference	no significant difference	no significant difference	-	-	-	-	-	
GRADE certainty of evidence	⊕⊕⊕◯ MODERATE[a]	-	⊕⊕◯◯ LOW[a,b]	-	-	-	-	-	-	
Judgement of effect (SMB)	**Large**	**trivial**	**small**	**small**	-	-	-	-	-	
Recurrence (1 study; high risk of bias)										
	ADM vs Placebo (95% CI) [study n]	**CBT vs Placebo (95% CI) [study n]**	**ADM vs CBT (95% CI) [study n]**	**ADM vs. ADM + CBT (95% CI) [study n]**	**CBT vs CBT + ADM (95% CI) [study n]**	**ADM vs TAU (95% CI) [study n]**	**CBT vs TAU (95% CI) [study n]**	**ADM vs WL (95% CI) [study n]**	**CBT vs. WL (95% CI) [study n]**	**Comments**
Continuation (13 – 24 weeks)	no studies	no studies	no studies	no studies	no studies	no studies	no studies	no studies	no studies	
	-	-	-	-	-	-	-	-	-	
Maintenance (≥ 25 weeks)	no studies	no studies	no studies	RR 0.99 (0.69 to 1.41) at 3 years* [1]	no studies	no studies	no studies	no studies	no studies	*those who recovered in acute phase were randomised to

	-	-	-	no significant difference	-	-	-	-	-	continuation or withdrawal of ADM
GRADE certainty of evidence	-	-	-	-	-	-	-	-	-	
Judgement of effect (SMB)	**-**	**-**	**-**	**no effect**	-	-	-	-	-	
Improvement in quality of life (6 studies, 4 continuation, 2 maintenance; 3 studies - high risk of bias, 3 studies – some concerns)										
	ADM vs Placebo (95% CI) [study n]	**CBT vs Placebo (95% CI) [study n]**	**ADM vs CBT (95% CI) [study n]**	**ADM vs. ADM + CBT (95% CI) [study n]**	**CBT vs CBT + ADM (95% CI) [study n]**	**ADM vs TAU (95% CI) [study n]**	**CBT vs TAU (95% CI) [study n]**	**ADM vs WL (95% CI) [study n]**	**CBT vs. WL (95% CI) [study n]**	**Comments**
Continuation (13 – 24 weeks)	no studies	no studies	total QOLIE-89 scores improved in both groups at 16 wks* ADM (SRT): 51±14.9 to 66.1±17.7; CBT:50.1 ±13.6 to 63.0 ±18.4 SMD 3.10 (-2.89 to 9.09) (trend favours ADM) [1]	no studies	no studies	significant improvement in WHO-5 wellbeing index with ADM (SRT) vs TAU#: 88.0 ± 0.26 vs. 64 ± 0.48 SMD 24.0 (23.88 to 24.12) (favours ADM) [1]	no studies	no studies	no studies	*patients with epilepsy and MDD #pateints with diabetes and MDD
	-	-	no significant difference	-	-	**significant**	-	-	-	
Maintenance (≥ 25 weeks)	no studies	no studies	no studies	no studies	no studies	no studies	no studies	no studies	no studies	
	-	-	-	-	-	-	-	-	-	
GRADE certainty of evidence	-	-	⊕◯◯◯[a,b,c] VERY LOW	-	-	⊕◯◯◯[a,b,c] VERY LOW	-	-	-	
Judgement of effect (SMB)	**-**	**-**	**no effect**	**-**	**-**	**large effect**	-	-	-	
Improvement in social functioning or activities of daily living (7 studies, 5 continuation, 2 maintenance; 2 studies – high risk of bias, 5 studies – some concerns)										
	ADM vs Placebo (95% CI) [study n]	**CBT vs Placebo (95% CI) [study n]**	**ADM vs CBT (95% CI) [study n]**	**ADM vs. ADM + CBT (95% CI) [study n]**	**CBT vs CBT + ADM (95% CI) [study n]**	**ADM vs TAU (95% CI) [study n]**	**CBT vs TAU (95% CI) [study n]**	**ADM vs WL (95% CI)**	**CBT vs. WL (95% CI) [study n]**	**Comments**

								[study n]		
Continuation (13 – 24 weeks)	mean differences in the total SDS score AGM vs placebo* at 24 weeks 7.25±1.14 points SMD -0.76 (-1.0 to -0.52) (favours ADM) [1]	no studies	change in WSAS at 6 months ADM: 23.3±9.6 to 9.9±9.4; CBT: 19.1±9.4 to 8.6±8.7 SMD 0.14 (-0.30 to 0.58) (trend favours CBT) [1]	change in WSAS at 6 months ADM: 23.3±9.6 to 9.9 ±9.4; ADM + CBT: 24.6 ± 7.0 to 11.1 ± 2.0 SMD -0.28 (-0.64 to 0.08) (trend favours ADM) [1]	change in WSAS at 6 months CBT: 19.1±9.4 to 8.6±8.7; ADM plus CBT: 24.6±7.0 to 11.1±2.0 SMD -0.60 (-1.05 to -0.16) (trend favours CBT) [1]	no studies	change in WSAS at 6 months CBT: 19.1±9.4 8.6±8.7; TAU: 21.9±11.5 to 10.3±10.3 SMD -0.18 (-0.68 to 0.33) (trend favours CBT) [1]	no studies	no studies	*only responders at 12 weeks entered continuation phase. Similar findings in severe MDD subgroup
	significant	-	no significant difference	no significant difference	no significant difference	-	no significant difference	-	-	
Maintenance (≥ 25 weeks)	no differences in changes of the HAMD work/activities item and SDS scores between ADM (DLX) vs placebo at 9 months significant improvement in SASS scores at month 9 from baseline in the ADM vs. placebo 13.81±1.04 vs 9.28±1.92 [1]	no studies	no studies	no studies	no studies	no studies	no studies	no studies	no studies	
	significant	-	-	-	-	-	-	-	-	
GRADE certainty of evidence	⊕⊕⊕◯[b] MODERATE	-	⊕◯◯◯[a,b,c] VERY LOW	⊕◯◯◯[a,b,c] VERY LOW	⊕◯◯◯[a,b,c] VERY LOW	-	⊕◯◯◯[a,b,c] VERY LOW	-	-	
Judgement of effect (SMB)	**large effect**	-	**don't know**	**don't know**	**don't know**	-	**don't know**	-	-	
SECONDARY OUTCOMES										
Response – proportion with ≥ 50% improvement **(17 studies, 14 continuation, 3 maintenance)**										

	ADM vs Placebo (95% CI) [study n]	**CBT vs Placebo (95% CI) [study n]**	**ADM vs CBT (95% CI) [study n]**	**ADM vs. ADM + CBT (95% CI) [study n]**	**CBT vs CBT + ADM (95% CI) [study n]**	**ADM vs TAU (95% CI) [study n]**	**CBT vs TAU (95% CI) [study n]**	**ADM vs WL (95% CI) [study n]**	**CBT vs. WL (95% CI) [study n]**	**Comments**
Continuation (13 – 24 weeks)	Study design I **RR 1.57 (0.94 to 2.62)** [3] Study design III **RR 1.65 (1.27 to 2.14)** [2] **NMA: 1.60 (1.08 to 2.36)**	**NMA: RR 2.13 (1.30 to 2.47)**	Pooled at 13 to 24 weeks **RR 0.79 (0.65 to 0.95)** (favours CBT) [3] NMA: RR 0.75 (0.52 to 1.08)	no studies	no studies	no studies	no studies	no studies	no studies	
	significant	**significant**	**pooled, significant** NMA, not significant	**-**	**-**	-	**-**	-	-	
Maintenance (≥ 25 weeks)	No studies	no studies	RR= 1.27 (0.63 to 2.55) [1]	RR 0.91 (0.82to 1.00) [1]	RR 1.40 (0.69 to 2.83) [1]	no studies	no studies	no studies	no studies	
	-	-	no significant difference	no significant difference	no significant difference	-	**-**	-		
GRADE certainty of evidence	-	-	-	-	-	-	-	-	-	
Judgement of effect (SMB)	**large effect**	**large effect**	**moderate effect**	**No effect**	-	-	-	-		
Remission (21 studies, 17 continuation, 4 maintenance)										
	ADM vs Placebo (95% CI) [study n]	**CBT vs Placebo (95% CI) [study n]**	**ADM vs CBT (95% CI) [study n]**	**ADM vs. ADM + CBT (95% CI) [study n]**	**CBT vs CBT + ADM (95% CI) [study n]**	**ADM vs TAU (95% CI) [study n]**	**CBT vs TAU (95% CI) [study n]**	**ADM vs WL (95% CI) [study n]**	**CBT vs. WL (95% CI) [study n]**	**Comments**
Continuation (13 – 24 weeks)	**RR 1.93 (1.38 to 2.69)** (favours ADM) [3] NMA **RR 1.98**	NMA **RR 3.01 (1.95 to 4.67)** (favours CBT) [1]	**RR 0.76 (0.59 to 0.98)** (favours CBT) [5] NMA: **RR 0.66**	**RR 0.40 (0.40 to 0.92)** (favours ADM + CBT) [2] NMA:	NMA RR 1.19 (0.75 to 1.93) [1]	NMA RR 0.71 (0.35 to 1.45) [1]	NMA RR 1.08 (0.51 to 2.28 [1]	no studies	no studies	

	(1.39 to 2.82) Favours ADM [3]		**(0.49 to 0.89)** Favours CBT [3]	**RR 0.55 (0.36 to 0.84)** Favours ADM + CBT [2]						
	significant	**significant**	**significant**	**significant**	no significant difference	no significant difference	no significant difference	-	-	
Maintenance (≥ 25 weeks)	no studies	no studies	0.69 (0.29 to 1.67) [1]	RR 0.95 (0.87 to 1.05) at 1 year [1]	RR 0.73 (0.3 to 1.76) indirect estimate [0]	no studies	no studies	no studies	no studies	
	-	-	no significant difference	no significant difference	no significant difference	-	-	-	-	
GRADE certainty of evidence	-	-	-	-	-	-	-	-	-	
Judgement of effect (SMB)	**moderate effect**	**moderate effect**	**moderate effect**	**moderate effect**	**no effect**	**no effect**	**no effect**	-	-	
SAFETY OUTCOMES										
Acceptability - proportion of participants who left the trial for any reason prior to the end of the study **(33 studies, 25 continuation, 9 maintenance)**										
	ADM vs Placebo (95% CI) [study n]	**CBT vs Placebo (95% CI) [study n]**	**ADM vs CBT (95% CI) [study n]**	**ADM vs. ADM + CBT (95% CI) [study n]**	**CBT vs CBT + ADM (95% CI) [study n]**	**ADM vs TAU (95% CI) [study n]**	**CBT vs TAU (95% CI) [study n]**	**ADM vs WL (95% CI) [study n]**	**CBT vs. WL (95% CI) [study n]**	**Comments**
Continuation (13 – 24 weeks)	RR 1.41 (0.93 to 2.14) [4] **RR= 1.69 (1.07 to 2.69)** (favours placebo) [4]* Study design III **RR 0.71 (0.57 to 0.89)** (favours ADM) NMA RR 1.23 (0.78 to 1.95) [4]	RR 0.65 (0.32 to 1.32) at 16 weeks NMA RR 0.88 (0.48 to 1.61) [1]	RR 1.57 (0.94 to 2.62) [6] NMA RR 1.41 (0.90 to 2.21) [5]	**RR= 2.06 (1.31 to 3.25)** at 24 weeks (favours ADM+CBT) [1] NMA **RR 2.45 (1.12 – 5.35)** (favours ADM+CBT) [1]	**RR 2.12, (1.29 to 3.48)** at 24 weeks, (favours ADM + CBT) [1] NMA RR 0.57 (0.26 to 1.27) [1]	RR 1.11 (0.78 to 1.60) [1] NMA RR 1.49 (0.69 to 3.22) [1]	RR 1.29 (0.80 to 2.08) at 24 weeks **[1]** NMA RR 1.06 (0.48 to 2.3)	no studies	no studies	*older participants
	No difference/ divergent	no significant difference	no significant difference	**significant**	**significant**	no significant difference	no significant difference	-	-	

Maintenance (≥ 25 weeks)	Study design I RR 0.96 (0.85 to 1.07) [3] Study design II RR 0.72 (0.4 to 1.29) [2] NMA RR 0.93 (0.82 – 1.05) [2]	no studies NMA RR 0.77 (0.54 to 1.08) indirect estimate [0]	RR 1.21 (0.87 to 1.68) at 96 weeks [1] **RR 0.62 (0.41 to 0.93)** at 44 weeks (favours ADM) [1]	RR 1.23 (0.92 to 1.65) [2] NMA RR 1.23 (0.93 to 1.63)	no studies NMA RR 0.98 (0.64 to 1.51) indirect estimate [0]	no studies	no studies	no studies	no studies	
	no significant difference	no significant difference	divergent	no significant difference	no significant difference	-		-	-	
GRADE certainty of evidence*	MODERATE[a]	LOW[a,b*]	LOW[a,b]	LOW[a,b*]	LOW[a,b*]	LOW[a,b*]	LOW[a,b*]	-	-	*for direct estimates
Judgement of effect (SMB)	**Don't know**									
Worsening of depression (2 studies, 2 continuation) risk of bias: some concerns										
	ADM vs Placebo (95% CI) [study n]	**CBT vs Placebo (95% CI) [study n]**	**ADM vs CBT (95% CI) [study n]**	**ADM vs. ADM + CBT (95% CI) [study n]**	**CBT vs CBT + ADM (95% CI) [study n]**	**ADM vs TAU (95% CI) [study n]**	**CBT vs TAU (95% CI) [study n]**	**ADM vs WL (95% CI) [study n]**	**CBT vs. WL (95% CI) [study n]**	**Comments**
Continuation (13 – 24 weeks)	no studies	no studies	RR 1.13 (0.36 to 3.54) at week 16 [1]	no studies	no studies	no studies	no studies	no studies	RR 0.86 (0.27 to 2.76) At week 16 [1]	
	-	-	no significant difference	-	-	-	-	-	no significant difference	
Maintenance (≥ 25 weeks)	no studies	no studies	no studies	no studies	no studies	no studies	no studies	no studies	no studies	
	-	-	-	-	-	-	-	-	-	
GRADE certainty of evidence	-	-	⊕⊕◯◯[a,c] LOW	-	-	-	-	-	⊕⊕⊕◯[b] MODERATE	
Judgement of effect (SMB)	**Don't know**									
Mortality (14 studies, 10 continuation, 4 maintenance) – 3 high risk of bias, all others had some concerns										

	ADM vs Placebo (95% CI) [study n]	CBT vs Placebo (95% CI) [study n]	ADM vs CBT (95% CI) [study n]	ADM vs. ADM + CBT (95% CI) [study n]	CBT vs CBT + ADM (95% CI) [study n]	ADM vs TAU (95% CI) [study n]	CBT vs TAU (95% CI) [study n]	ADM vs WL (95% CI) [study n]	CBT vs. WL (95% CI) [study n]	Comments
Continuation (13 – 24 weeks)	No deaths at 24 weeks [2] 1 possible related cardiac death [1}	no studies	1 unrelated death at 16 weeks [1]	no studies	no studies	no studies	no studies	no studies	no studies	
	-	-	-	-	-	-	-	-	-	
Maintenance (≥ 25 weeks)	10 vs 5 deaths with ADM vs. placebo by week 39 [1]	no studies	No studies	no studies	no studies	no studies	no studies	no studies	no studies	
		-	-	-	-	-	-	-	-	
GRADE certainty of evidence	⊕⊕◯◯[d] LOW	-	⊕⊕◯◯[b,c] LOW	-	-	-	-	-	-	
Judgement of effect (SMB)	**Don't know**									

Legend: ADM: antidepressant medication; CBT: cognitive behavioural therapy; TAU: treatment as usual; WL: waiting list; RR: risk ratio; SMD: standardized mean difference; CI: confidence interval; RCT: randomized controlled trial; AGM – agomelatine; FXT - fluoxetine; WSAS, WHO – world health organization, HAMD - Hamilton Rating Scale of Depression, SDS- Sheehan Disability Scale, SASS- Social Adaptation Self-evaluation Scale, NMA – network meta-analysis

GRADE Working Group grades of evidence (summarized from Tables 4 and 9 in the Assessment Report)

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

a. Downgraded one point as majority of studies judged as of overall poor quality regarding risk of bias; b. Downgraded one point due to imprecision (defined as wide confidence intervals including no effect or low overall sample size (defined as <400 participants for continuous outcomes or below optimal information size for dichotomous outcomes)); c. Downgraded one point due to indirectness related to the population of interest; d. Downgraded two points due to indirectness related to the population of interest and the different designs of the RCTs.