BMJ Open

Beneficial and harmful effects of duloxetine versus placebo, 'active placebo' or no intervention for adults with major depressive disorder: a systematic review with meta-analysis and trial sequential analysis of randomised clinical trials

Faiza Siddiqui ,[1] Johanne Juul Petersen ,[1] Sophie Juul ,[1,2,3] Caroline Barkholt Kamp ,[1,4] Marija Barbateskovic ,[1] Joanna Moncrieff ,[5,6] Mark Abie Horowitz ,[6,7] Mathias Maagaard,[8] Kiran Kumar Katakam,[1] Christian Gluud ,[1,4] Janus C Jakobsen[1,4]

To cite: Siddiqui F, Petersen JJ, Juul S, *et al*. Beneficial and harmful effects of duloxetine versus placebo, 'active placebo' or no intervention for adults with major depressive disorder: a systematic review with meta-analysis and trial sequential analysis of randomised clinical trials. *BMJ Open* 2025;**15**:e082853. doi:10.1136/bmjopen-2023-082853

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2023-082853).

Received 05 December 2023
Accepted 08 January 2025

For numbered affiliations see end of article.

Correspondence to
Dr Faiza Siddiqui;
faiza.siddiqui@ctu.dk

ABSTRACT

Objectives To assess the beneficial and harmful effects of duloxetine versus 'active placebo', placebo or no intervention for adults with major depressive disorder.

Design Systematic review with meta-analysis and trial sequential analysis of randomised trials.

Data sources Cochrane Central Register of Controlled Trials, MEDLINE, Embase, PsycINFO and other relevant databases up until January 2023. We requested clinical study reports from 36 competent authorities.

Eligibility criteria for selecting studies All randomised clinical trials comparing duloxetine versus placebo, 'active placebo' or no intervention, irrespective of publication type, publication status, publication year and language for treatment of major depressive disorder in adults.

Data extraction and synthesis Five authors in pairs extracted data using a standardised data extraction sheet. A third review author was consulted for disagreements. Intervention effects were assessed by both random-effects and fixed-effect model meta-analyses, risk of bias assessments were performed by two independent review authors using Cochrane's risk of bias tool V.2 and the certainty of evidence was assessed using Grading of Recommendations Assessment, Development and Evaluation.

Results We included 28 trials randomising a total of 7872 participants. All results were at high risk of bias. The trials' assessment time points were between 6 and 16 weeks after randomisation. Meta-analyses showed evidence of a beneficial effect of duloxetine on depressive symptoms (mean difference −1.81, Hamilton Depression Rating Scale (HDRS-17) points; 95% CI −2.34 to −1.28; heterogeneity I^2=0.0%; 12 trials) and quality of life (mean difference −3.79 points, 95% CI −5.11 to −2.46; I^2=0.0%; three trials), but the effect sizes were below our predefined minimal clinically important differences. Trial sequential analysis showed that we did not have enough information to assess the effects of duloxetine on serious adverse events (SAEs) (OR 0.67, 95% CI 0.44 to 1.02; I^2=0.0%; 19 trials) or suicide or suicide attempts (OR 1.08, 95% CI 0.37 to 3.16; six trials). Duloxetine increased the risk of non-SAEs (risk ratio 1.27, 95% CI 1.22 to 1.32; I^2=73.0%; 24 trials). The adverse events with the lowest number needed to harm (NNH) were nausea (NNH 6), dry mouth (NNH 13), somnolence (NNH 17), withdrawal syndrome (NNH 19), sweating (NNH 20), dizziness (NNH 21) and constipation (NNH 21).

Conclusions Duloxetine appears to reduce depressive symptom scores and improve quality of life scores in the short term, but the effect sizes are minimal and of questionable patient importance. The short- and long-term effects of duloxetine on risks of SAEs and suicidality are uncertain. Duloxetine increases the risks of several short-term adverse events. Systematic assessments of benefits and harms over longer periods are required.

Trial registration number PROSPERO 2016 CRD42016053931.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ The present review took into account the risks of systematic errors, random errors, generalisability, publication bias and heterogeneity.
⇒ The current evidence on the effects of duloxetine for major depressive disorder is based on trials at high risk of bias, which leads to risks of overestimating the beneficial effects of duloxetine.
⇒ Only the short-term (6 to 16 weeks after randomisation) effects of duloxetine are known.

INTRODUCTION

The diagnosis of major depressive disorder is based on symptoms such as persistent low mood and loss of interest in normally

enjoyable activities, decreased appetite, sleep disturbances and occasionally suicidal thoughts and suicide attempts.[1] It is one of the leading causes of disability, impaired quality of life and decreased work productivity worldwide.[2–6] Serotonin-norepinephrine reuptake inhibitors (SNRIs) constitute a relatively new class of antidepressants.[7] The SNRI duloxetine is widely prescribed for the treatment of depression in the USA and Europe.[8–10] Previous evidence has shown that duloxetine on average reduces depressive symptoms with a statistically significant effect, but the clinical importance of this effect is uncertain.[11 12] Moreover, previous assessments of the adverse effects of duloxetine have been inadequate. Some studies have found that antidepressants increase the risks of suicide and suicide attempts, particularly during the initiation of treatment and also during ongoing treatment, but this has not been examined specifically for duloxetine.[13]

Our objective was to assess the beneficial and harmful effects of duloxetine versus placebo, 'active placebo' or no intervention in the treatment of adults with major depressive disorder.

METHODS

The methodology for this review as well as inclusion and exclusion criteria for the studies is described in detail in our published protocol.[14] This review is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[15] In short, we included all randomised clinical trials comparing duloxetine versus placebo, 'active placebo' or no intervention irrespective of publication type, publication status, publication year and language for the treatment of major depressive disorder in adults.

Outcomes

Primary outcomes

- Depressive symptoms measured with the Hamilton Depression Rating Scale (HDRS-17).[16] Minimal important difference on HDRS-17 was predefined as 3 points. Although there is no consensus on this issue, we chose a difference of 3 points on HDRS-17 (or 0.5 standardised mean difference) as minimal important difference, based on National Institute of Health and Clinical Excellence (NICE) guidelines on treating depression.[17] This is supported by more recent work by Hengartner and Ploderl where they used both within-patient and between-patient anchor-based approaches and concluded that minimal important difference on HDRS-17 is likely to be in the range of 3 to 5 points.[14 17 18]
- Serious adverse events. We used the International Conference on Harmonisation of technical requirements for registration of pharmaceuticals for human use-good clinical practice (ICH-GCP) definition of a SAE.[19]

Secondary outcomes

- Suicides or suicide attempts (as defined by the trialists).
- Quality of life (assessed with any valid continuous quality of life questionnaire such as Quality of Life in Depression Scale (QLDS), EQ-5D or any other scale used by the trialists).
- Suicidal ideation (as defined by the trialists).

Exploratory outcomes

- Non-SAEs (any adverse event not defined as serious).
- Depressive symptoms measured using any form of HDRS, Montgomery-Asberg Depression Rating Scale or Beck's Depression Inventory.
- Response to treatment (defined as a 50% reduction from baseline on any scale).
- Remission (as defined by trialists).

Outcomes were assessed at the end of treatment and at maximum follow-up.[14]

Search methods

Cochrane Central Register of Controlled Trials, MEDLINE, Embase, PsycINFO, Science Citation Index Expanded, Social Sciences Citation Index, Conference Proceedings Citation Index-Science and Conference Proceedings Citation Index-Social Science & Humanities, Chinese databases (CNKI, Wanfang, VIP, Sinomed) and Google Scholar were searched from their inception until 23 January 2023 (Supplement I: Search strategy). We also searched trial registers of pharmaceutical companies, the WHO trial registry, ClinicalTrials.gov, including the websites of the Food and Drug Administration (FDA) and the European Medicines Agency (EMA). Furthermore, we requested clinical study reports from 36 competent authorities including the FDA and the EMA.

Inclusion and exclusion criteria

We included all randomised clinical trials comparing duloxetine at any dose or duration with placebo, 'active placebo' or no intervention for major depressive disorders in adults irrespective of publication type, publication status, publication year and language. We excluded trials exclusively including participants with a somatic disease and comorbid major depressive disorder and trials on major depressive disorder during or after pregnancy.

Screening and data extraction

Five authors in pairs (FS, MB, SJ, CBK and JJP) extracted data using a standardised data extraction sheet. We extracted data pertaining to trial characteristics (follow-up period, funding and number randomised), participant characteristics (age, baseline depression scores and chronic or treatment-resistant depression) and intervention (duration, dose, placebo wash out and cointerventions) and outcomes. A third review author (JCJ) was consulted for disagreements. Risk of bias assessments were performed by two independent review authors using Cochrane's risk of bias tool V.2.[20]

Statistical methods

Intervention effects were assessed by both random-effects (Sidik-Jonkman tau-estimator)[21] and fixed-effect model (Mantel-Haenszel) meta-analyses using Stata V.17[14] and RStudio.[22] We primarily reported the most conservative point estimate (the highest p value) of the two and considered the less conservative result as a sensitivity analysis.[23 24] We assessed a total of five primary and secondary outcomes; therefore, we considered p≤0.016 as statistically significant.[24] We used an eight-step procedure to assess if the thresholds for significance were crossed.[24] We adjusted for zero-event cells using treatment-arm continuity correction.[23 25] For trials with multiple relevant experimental groups, we divided the number of events and sample size of the control group for dichotomous outcomes and divided the sample size and kept the mean and SD of the control group for continuous outcomes. If the data could not be equally divided due to an odd number of events, we drew lots to decide which comparison would be favoured. When necessary, we divided the control group to provide for subgroup analyses. For continuous outcomes, we prioritised end scores but included changes from baseline scores if only these were reported.[23]

We performed trial sequential analyses on all the outcomes.[26–28] In trial sequential analysis for dichotomous outcomes, we used the proportion of participants with an outcome in the control group and a relative risk reduction of 25%. For continuous outcomes, we used the empirical SD, a mean difference of three points on the HDRS (17 or 21 items) and the observed SD/2 when other depression scales or quality of life scales are used. For all outcomes, we used an alpha of 1.7%, a beta of 10% and adjustment for the observed diversity of the trials in the meta-analysis. Heterogeneity was assessed using forest plots and I^2 statistics. We also planned several subgroup analyses based on risk of bias in trials, risk of for-profit bias, type of control intervention, participants' age (≥65 years or <65 years), placebo washout period, duration of intervention (<6 weeks, 6–12 weeks and >12 weeks), baseline HDRS scores (<23 or ≥23), chronic or treatment-resistant depression (excluded/not excluded) and duloxetine dose (≤60 mg/day or >60 mg/day).[14]

Grading of Recommendations Assessment, Development and Evaluation

We assessed the certainty of evidence of primary and secondary outcomes using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) tool.[23] We used the five GRADE domains (bias risk of the trials, consistency of effect, imprecision, indirectness and publication bias) to assess the quality of evidence. Imprecision was assessed using trial sequential analysis. We downgraded imprecision in GRADE by two levels if the accrued number of participants was below 50% of the diversity-adjusted required information size (DARIS), and one level if between 50% and 100% of DARIS.

Patient and public involvement

A person with lived experience of taking antidepressants for depression was included in the research group to feed into decision-making. During the protocol development, we chose outcomes that we believe are important from patients' perspectives.

Results

We identified 28 randomised clinical trials, 20 published articles,[29–48] 2 clinical study summaries available online (11918A, H8I-MC-HQAC),[49 50] 1 trial registry (NCT 01145755)[47] and 5 clinical study reports (study IDs; F1J-MC-HMAI,[51] F1J-MC-HMAQ,[52] F1J-MC-HMAT,[53] F1J-MC-HMAG[54] and F1J-MC-HMAH)[55] from the EMA comparing duloxetine versus inert placebo for the treatment of major depressive disorder (online supplemental table S1 and figure S1). We also retrieved clinical study reports from the EMA corresponding to six included trials with published results.[32–36 43] We were unable to identify any trials assessing duloxetine versus 'active placebo' or no intervention.

In total, 4562 participants were randomised to duloxetine versus 3310 randomised to placebo. All included trials were at high risk of bias (figure 1) and were funded by industry except one for which we had no information on funding.[48] All trials assessed outcomes at the end of treatment period (6 to 16 weeks), and no long-term follow-up data were available. The proportion of participants with missing data at follow-up was not adequately reported in most trials; therefore, it was not possible to perform 'best-worst/worst-best' sensitivity analyses.

Primary outcomes

Hamilton Depression Rating Scale-17

12 trials reported mean HDRS-17 follow-up scores or change scores and corresponding SD at the end of treatment.[32–36 43 48 51–55] All trials only assessed outcomes at the end of the treatment period, that is, from 6 to 12 weeks after randomisation. Meta-analysis showed that duloxetine versus placebo reduced depressive symptoms (mean difference −1.81 points, 95% CI −2.34 to −1.28; p<0.01; I^2=0.0%; 12 trials; Bayes factor 2.88×10^{-6}) (online supplemental figure S2). However, the effect size −1.81 was below our predefined minimal important difference −3.0. Trial sequential analysis showed that the boundary for benefit was crossed confirming the statistically significant meta-analysis result (online supplemental figure S3). This outcome result was assessed at high risk of bias, and the certainty of the evidence was low (online supplemental figure S4 and figures 1 and 2).

None of the following subgroup analyses showed evidence of a statistically significant difference: (1) trials with mean baseline HDRS scores <23 points (n=11 trials) compared with one trial with mean baseline HDRS scores ≥23 points (p=0.28; online supplemental figure S5); (2) trials excluding participants with chronic or treatment-resistant depression (n=9 trials) compared with trials not excluding participants with chronic or treatment-resistant

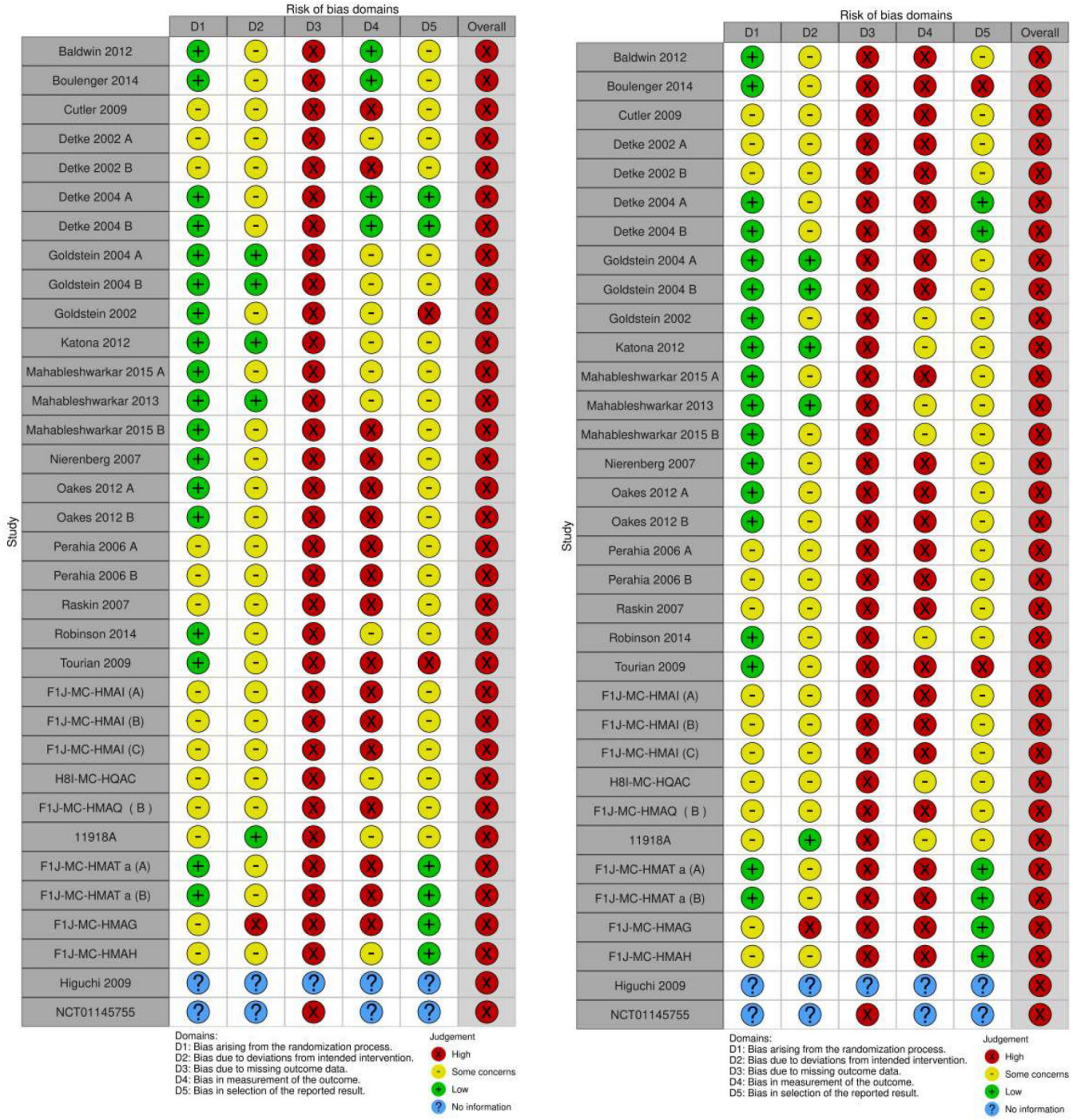


Figure 1 Risk of bias in the included trials for efficacy (left) and safety (right) outcomes.

depression (n=2 trials) (p=0.19; online supplemental figure S6); (3) trials with duloxetine dose ≤60 mg/day (n=8 trials) compared with >60 mg/day (n=4 trials) compared with variable dose (n=2 trials) (p=0.99; online supplemental figure S7) and (4) trials without placebo washout period (n=11 trials) compared with one trial where some participants had placebo washout period (p=0.73; online supplemental figure S8). The remaining subgroup analyses were not possible to conduct due to lack of relevant data.

Serious adverse events

23 trials reported the proportion of participants with SAEs.[29 30 32–44 46 47 49–53 55] SAEs were reported either at the end of treatment or up to 4 weeks after the end of study treatment (6 to 16 weeks). A total of 40 of 3948 duloxetine participants (1.01%) experienced SAE compared with 45 of 2816 placebo participants (1.59%). Meta-analysis showed no evidence of a difference in occurrence of SAE (OR 0.67, 95% CI 0.44 to 1.02; p=0.06; I^2=0.0%; 19 trials) (figure 3). Binomial regression showed comparable

Summary of findings:

Duloxetine compared to placebo for major depressive disorders

Patient or population: major depressive disorder

Setting:

Intervention: [intervention]

Comparison: [comparison]

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with placebo	Risk with duloxetine				
Hamilton Depression Rating Scale-17 (HDRS-17) Scale from: 0 to 52	The mean hamilton Depression Rating Scale-17 was **0** points	MD **1.81 points lower** (2.34 lower to 1.28 lower)	-	2950 (12 RCTs)	⨁⨁◯◯ Low[a]	
Serious adverse events (SAE)	16 per 1,000	**11 per 1,000** (7 to 16)	**RR 0.67** (0.44 to 1.02)	6764 (23 RCTs)	⨁◯◯◯ Very low[a,b]	
Suicide or suicide attempt	1 per 1,000	**1 per 1,000** (0 to 4)	**RR 1.23** (0.43 to 3.53)	6815 (24 RCTs)	⨁◯◯◯ Very low[a,b,c]	
Quality of life	The mean quality of life was **0**	MD **3.79 lower** (5.11 lower to 2.46 lower)	-	660 (3 RCTs)	⨁⨁◯◯ Low[a]	
Suicide ideation	29 per 1,000	**27 per 1,000** (11 to 66)	**RR 0.94** (0.39 to 2.27)	1764 (6 RCTs)	⨁◯◯◯ Very low[a,b,d]	
Non-serious adverse events (NSAE)	527 per 1,000	**669 per 1,000** (643 to 695)	**RR 1.27** (1.22 to 1.32)	7002 (24 RCTs)	⨁◯◯◯ Very low[a,e]	

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: confidence interval; **MD:** mean difference; **RR:** risk ratio

GRADE Working Group grades of evidence
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.

Explanations

a. a. Downgraded 2 for risk of bias. All trials at high risk of bias in atleast one domain and with for-profit bias.

b. b. Downgraded 2 for imprecision. Trial Sequential Analysis showing that there was not enough information to confirm or reject a relative risk reduction (RRR) of 25% and the accrued number of participants is below 50% of the diversity-adjusted required information size (DARIS). Also, very few events in each trial and CI overlaps no effect

c. In many trials, suicide was not reported specifically so we regarded no deaths reported as no suicide

d. moderate heterogeneity

e. Downgraded 1 for substantial heterogeneity

Figure 2 Summary of findings table. GRADE, Grading of Recommendations Assessment, Development and Evaluation.

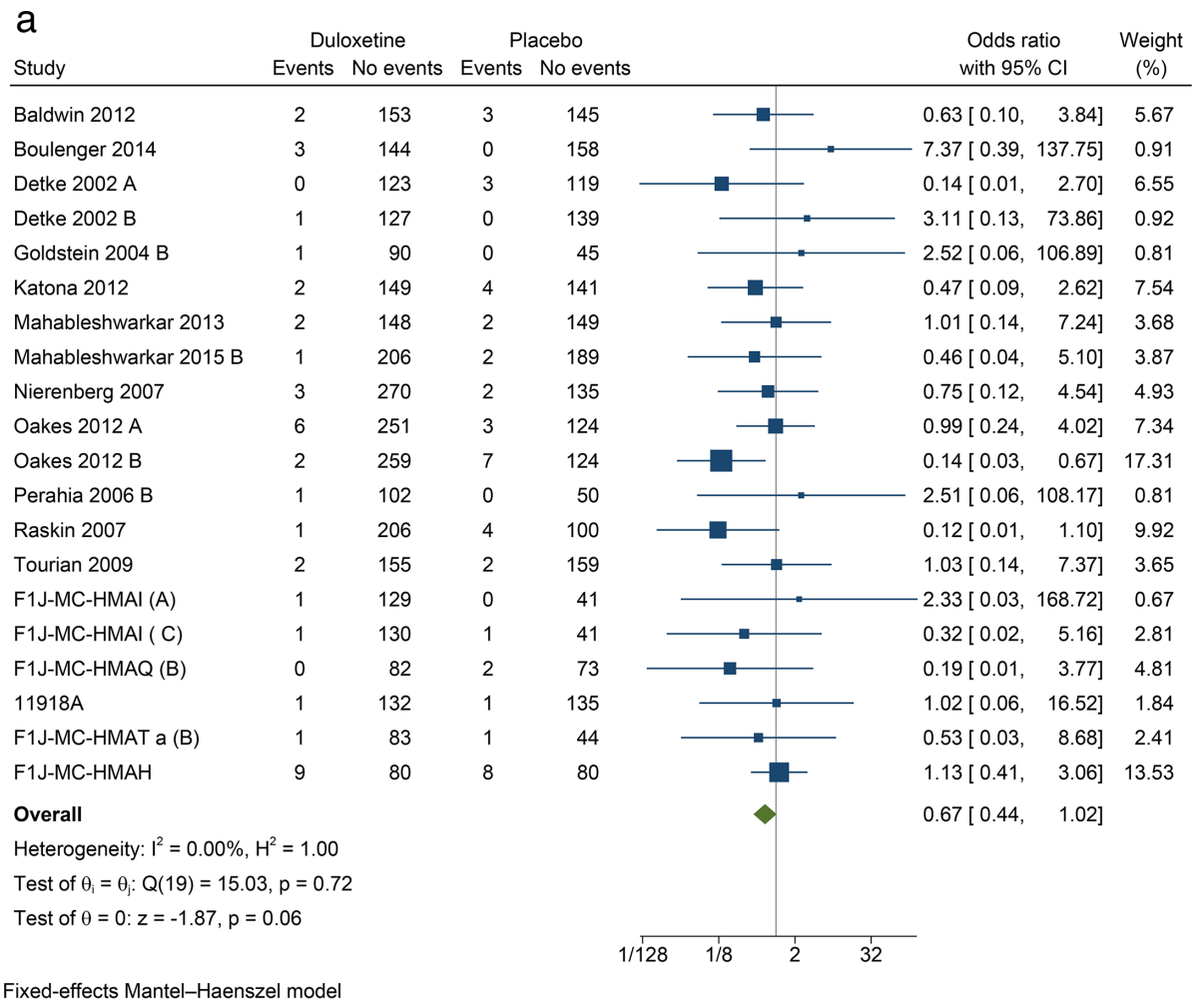


Figure 3 Meta-analysis (A) and trial sequential analysis (B) of duloxetine versus placebo on serious adverse events.

results (OR 0.64, 95% CI 0.41 to 1.00; p=0.05). Online supplemental table S2 summarises the specific SAE in the included trials. Trial sequential analysis showed that we did not have enough information to confirm or reject the hypothesis that duloxetine decreased the risk of SAE with a relative risk reduction of 25% (figure 3). This outcome

result was assessed at high risk of bias, and the certainty of the evidence was very low (figure 1, online supplemental figures S2 and S9).

Subgroup analysis comparing the effects of baseline HDRS scores (p=0.63), chronic or treatment-resistant depression (p=0.96), participants' age (p=0.15), duloxetine dose (p=0.54) and placebo washout period (p=1.00) showed no evidence of a difference (online supplemental figures S10–S14). We meta-analysed four individual SAEs, namely, accidental injury (p=0.57), myocardial infarction (p=0.99), depression (p=0.38) and asthma (p=0.27); however, no significant differences were observed.

Secondary outcomes

Suicide or suicide attempt

24 trials reported on suicide and suicide attempts.[29 30 32–46 49–52 54 55] All trials assessed outcomes either at the end of treatment or up to 4 weeks after the end of study treatment (6 to 16 weeks). 7 of 1683 duloxetine participants (0.4%) attempted suicide versus 2 of 1165 placebo participants (0.2%). Meta-analysis showed no evidence of a difference in suicide or suicide attempts (OR 1.23, 95% CI 0.43 to 3.53; p=0.69; I^2=0.0%; six trials) (figure 4). Binomial regression showed comparable results (OR 1.43, 95% CI 0.35 to 5.75; p=0.61). Trial sequential analysis showed that we did not have enough information to confirm or reject that duloxetine reduced the risk of suicide or suicide attempt by 25% (figure 4). This outcome result was assessed at a high risk of bias, and the certainty of the evidence was very low (figure 1, online supplemental figures S2 and S15).

Subgroup analyses comparing the effects of baseline HDRS scores (p=0.49), chronic or treatment-resistant depression (p=0.38), participants' age (p=0.55) and placebo washout period (p=0.56) showed no evidence of a difference (online supplemental figures S16–S19).

Quality of life

Six trials reported data on quality of life.[30 32 35 45 52 53] Three out of six trials reported mean change in quality of life scores and corresponding SD using QLDS.[32 35 53] Meta-analysis of these three trials showed evidence of a beneficial effect of duloxetine on quality of life (mean difference −3.79 points, 95% CI, −5.11 to −2.46; p<0.001; I^2=0.0%; three trials; Bayes factor $1.88*10^{-7}$) (online supplemental figure S20), but the effect was below our predefined threshold of 4.14 points. Trial sequential analysis showed that the boundary for benefit was crossed confirming the statistically significant meta-analysis result (online supplemental figure S21). This outcome was assessed at a high risk of bias, and the certainty of the evidence was low. Subgroup analysis comparing the effects of duloxetine dose showed no evidence of a difference (p=0.83; online supplemental figure S22). This outcome was assessed at a high risk of bias and the certainty of the evidence was low.

Suicidal ideation

Six trials reported data on suicidal ideation.[29 30 37–39 46] Suicidal ideation was reported by 20 of 873 duloxetine participants (2.3%) compared with 26 of 891 placebo participants (2.9%). Meta-analysis showed no evidence of a difference in suicidal ideation (risk ratio (RR) 0.94, 95% CI 0.39 to 2.27; p=0.36; six trials) (online supplemental figure S23). Visual assessment of the forest plot and statistical tests (I^2=31.4%) showed moderate heterogeneity. It was not possible to perform trial sequential analysis due to too little information. This outcome result was assessed at a high risk of bias and the certainty of the evidence was very low (figures 1,2).

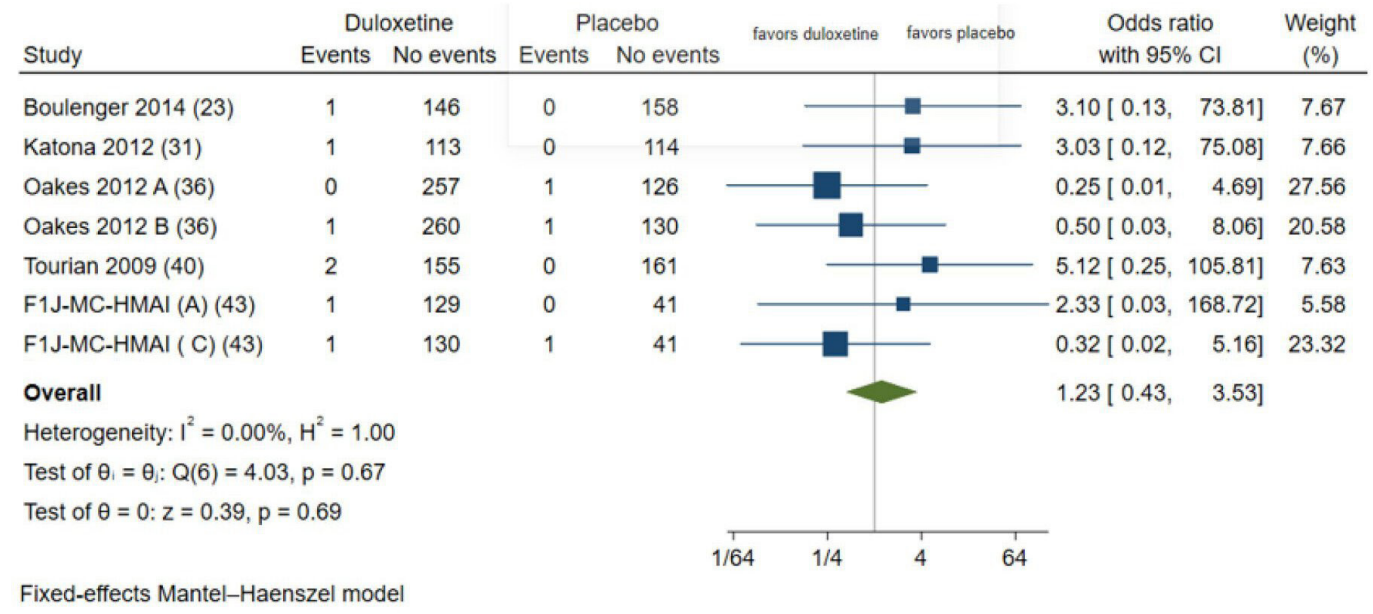
Exploratory outcomes

Non-serious adverse events

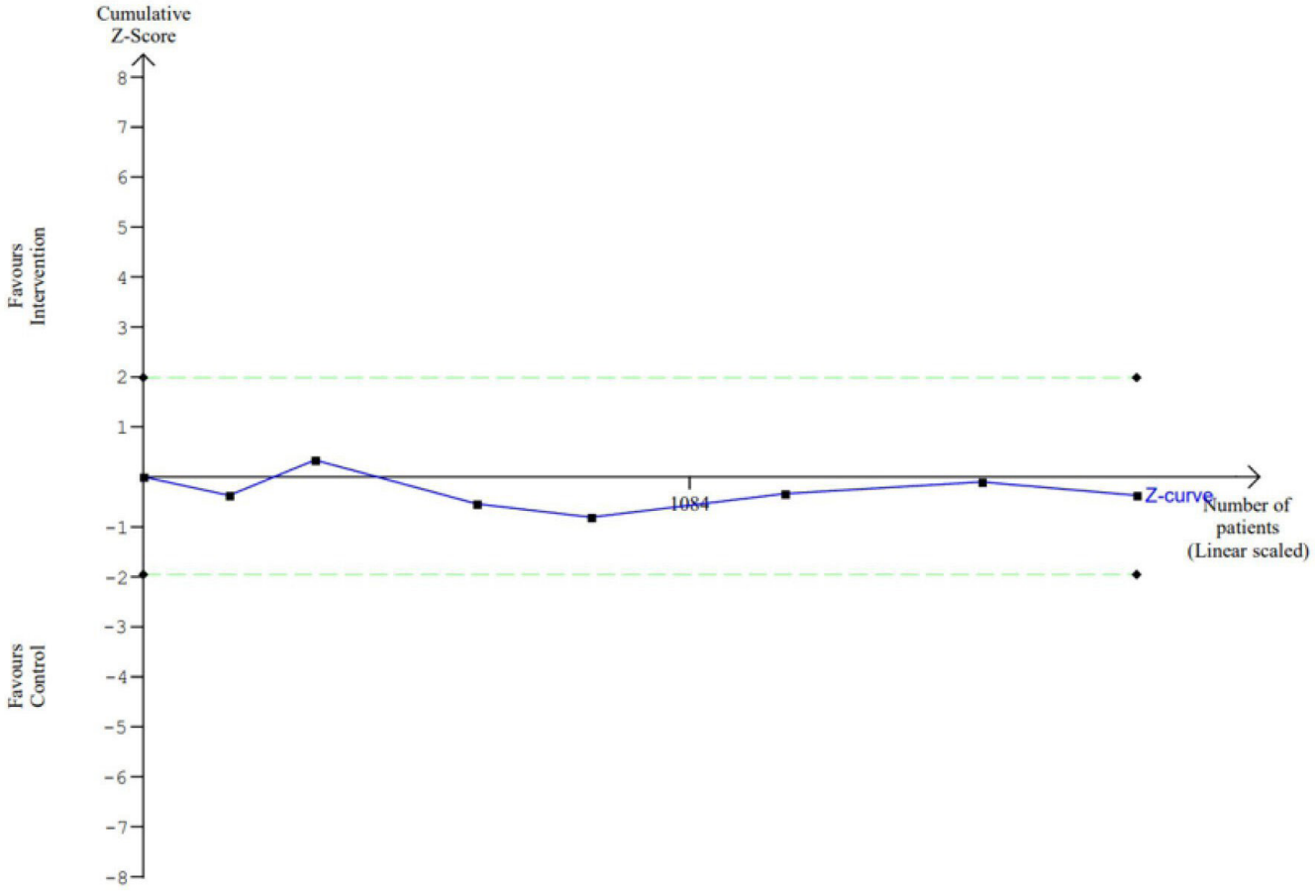
Data on non-SAEs were reported by 24 trials.[29 30 32–44 46 47 49 51–53 55] A total of 2625 of 4061 duloxetine participants experienced non-SAEs (64.6%) versus 1549 of 2941 placebo participants (52.7%). Non-SAEs were reported either at the end of treatment or up to 4 weeks after the end of study treatment (6 to 16 weeks). Meta-analysis showed that duloxetine versus placebo increased the risk of overall non-SAEs compared with placebo (RR 1.27, 95% CI 1.22 to 1.32; p<0.01; I^2=73.0%; 24 trials; Bayes factor 0.11) (figure 5). Trial sequential analysis showed that the boundary for harm was crossed confirming the meta-analysis result (figure 5). This outcome was assessed at a high risk of bias, and the certainty of evidence was very low (figures 1,2). Visual inspection of the funnel plot showed signs of asymmetry (online supplemental figure S24).

Subgroup analyses comparing the effects of baseline HDRS scores (p=0.37), participants' age (p=0.19) and duloxetine dose (p=0.13) showed no evidence of a difference (online supplemental figures S25–S27).

Subgroup analysis comparing 19 trials that excluded participants with chronic or treatment-resistant depression to the five trials that did not exclude participants with chronic or treatment-resistant depression showed evidence of a difference (p<0.01; online supplemental figure S28). When the subgroup of trials that excluded participants with chronic or treatment-resistant depression was analysed separately, meta-analysis showed evidence of a difference (RR 1.32, 95% CI 1.26 to 1.38; p<0.01; 19 trials) (online supplemental figure S29). Meta-analysis of trials that did not exclude participants with chronic or treatment-resistant depression also showed evidence of a difference with slightly lower relative risk (RR 1.15, 95% CI 1.07 to 1.22; p<0.01; five trials) (online supplemental figure S30). Subgroup analysis comparing eight trials reporting no placebo washout period compared with 15 trials reporting placebo washout period showed evidence of a difference (p<0.01; online supplemental figure S31). Meta-analysis of trials that reported no placebo washout period showed evidence of a difference (RR 1.40, 95% CI 1.30 to 1.50; p<0.01; eight trials) (online supplemental figure S32). Meta-analysis of

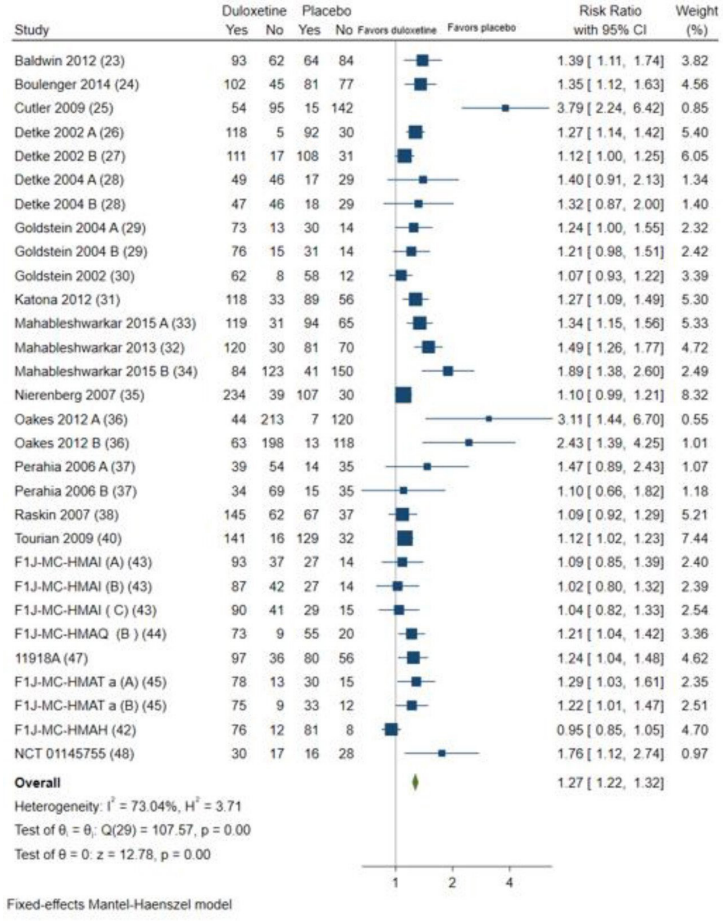


A *Meta-analysis*

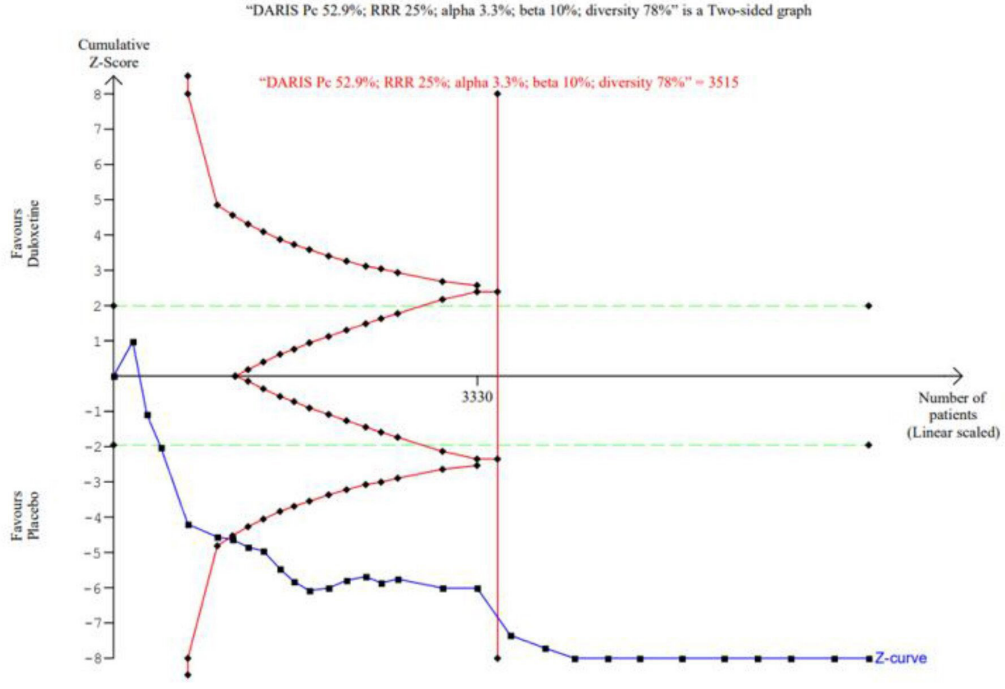


B *Trial Sequential Analysis*

Figure 4 Meta-analysis (A) and trial sequential analysis (B) of duloxetine versus placebo on suicide and suicide attempts.



A *Meta-analysis*



B *Trial Sequential Analysis*

Figure 5 Meta-analysis (A) and trial sequential analysis (B) of duloxetine versus placebo on non-serious adverse events.

trials that reported placebo washout period also showed evidence of a difference (RR 1.21, 95% CI 1.16 to 1.26; p<0.01; 15 trials) (online supplemental figure S33). The remaining predefined subgroup analyses were not possible to perform due to lack of relevant data.

We identified a total of 224 individual non-SAEs. We meta-analysed each specific non-SAE separately.

16 meta-analyses showed evidence of a harmful effect of duloxetine on individual non-SAEs (online supplemental table S3). The 16 adverse events ranked on the basis of NNH were nausea (RR 2.92, 95% CI 2.38 to 3.58; p<0.0001; 30 trials; NNH 6), dry mouth (RR 2.05, 95% CI 1.67 to 2.52; p<0.0001; 30 trials; NNH 12), somnolence (RR 2.40, 95% CI 1.81 to 3.18; p<0.0001; 26 trials; NNH 17), withdrawal syndrome defined as any events that were deemed by the investigator to be caused by cessation of medicine (RR 2.09, 95% CI 1.35 to 3.24; p<0.0009; one trials; NNH 19), sweating (RR 2.88, 95% CI 1.95 to 4.26; p<0.0001; 27 trials; NNH 20), dizziness (RR 1.88, 95% CI 1.49 to 2.38; p<0.0001; 30 trials; NNH 21), constipation (RR 1.79, 95% CI 1.30 to 2.47; p<0.0004; 28 trials; NNH 21), decreased appetite (RR 3.33, 95% CI 1.80 to 6.15; p<0.0001; nine trials; NNH 24), anorexia (RR 2.85, 95% CI 1.60 to 5.06; p<0.0004; 15 trials; NNH 25), insomnia (RR 1.64, 95% CI 1.27 to 2.11; p<0.0001; 25 trials; NNH 29), fatigue (RR 2.06, 95% CI 1.28 to 3.3; p<0.0029; 13 trials; NNH 30), vomiting (RR 2.28, 95% CI 1.4 to 3.71; p=0.0009; 22 trials; NNH 32), yawning (RR 5.54, 95% CI 2.32 to 13.22; p<0.0001; 11 trials; NNH 35), vasodilatation (RR 2.08, 95% CI 1.11 to 3.89; p<0.0215; 13 trials; NNH 41), diarrhoea (RR 1.38, 95% CI 1.12 to 1.71; p<0.0028; 30 trials; NNH 42) and decreased libido (RR 2.37, 95% CI 1.22 to 4.61; p<0.0111; 13 trials; NNH 50).

Duloxetine had a protective effect on some non-SAEs, namely, back pain (RR 0.64, 95% CI 0.42 to 0.95; p=0.028; 20 trials), hyperventilation (RR 0.14, 95% CI 0.02 to 0.82; p=0.029; three trials), pain (RR 0.65, 95% CI 0.43 to 0.97; p=0.03; 16 trials) and breast pain (RR 0.33, 95% CI 0.11 to 0.99; p=0.048; eight trials). The remaining exploratory outcomes are reported in the supplementary material (online supplemental figures S34 and S35).

DISCUSSION

In this systematic review, we aimed to assess the beneficial and harmful effects of duloxetine for adults diagnosed with major depressive disorder. We analysed 28 placebo-controlled clinical trials randomising a total of 7872 participants. Meta-analysis and trial sequential analysis showed that duloxetine versus placebo reduces depressive symptoms and increases quality of life with a statistically significant effect, but the effect sizes were below our minimal important differences. Meta-analysis and trial sequential analysis showed that we did not have enough information to confirm or reject the effects of duloxetine on SAEs and suicides or suicide attempts. We observed an increased risk of non-SAEs in participants receiving duloxetine compared with placebo. The ten adverse events with the lowest NNH were nausea, dry mouth, withdrawal syndrome, dizziness, constipation, somnolence, insomnia, diarrhoea, excessive sweating and fatigue. We were unable to identify any trials assessing duloxetine versus 'active placebo'.

Our meta-analysis confirmed the small statistically significant effect of duloxetine on depressive symptoms (−1.81 points), as observed in earlier systematic reviews, which was below our predefined minimal important difference (−3.0 points). However, it may be even less clinically significant than this since it has been argued that a difference of 7 to 8 points on HDRS is necessary to observe a clinical effect and factors like unblinding due to adverse effects, short treatment duration and publication bias tend to exaggerate treatment effects.[56] The small beneficial effects of duloxetine on depressive symptoms observed in this meta-analyses are in line with earlier meta-analysis of antidepressants, thereby highlighting the need to weigh these minimal beneficial effects in relation to adverse effects in clinical settings.[11 12] Moreover, HDRS has been criticised as an outcome to measure depressive symptoms due to its psychometric flaws.[57] We also observed a small beneficial effect on quality of life; however, the effect size was also below our predefined threshold. The minimal important difference in quality of life questionnaires has not been quantified adequately and quality of life is strongly affected by selective reporting.[58]

Most systematic reviews published so far have only focused on beneficial effects, and assessment of safety has been limited to proxy measures like tolerability or drop-outs owing to the adverse effects. This was also the case for the most recent systematic review and network meta-analysis of 21 antidepressants,[12] which assessed acceptability for antidepressants rather than specifically reported SAEs and non-SAEs. We extracted and meta-analysed all individual non-SAEs. We were able to retrieve clinical study reports from the EMA that provided us access to data on more than 200 non-SAEs that are not reported in published articles and trial registries. To the best of our knowledge, this has not been done before.

Other strengths of our review include predefined methodology, which was based on PRISMA guidelines, trial sequential analysis, the eight-step procedure defined by Jakobsen *et al*[24] and the GRADE approach. Thereby, we accounted for the risks of systematic errors, random errors, generalisability, publication bias and heterogeneity.

One of the major limitations of our review is the lack of long-term follow-up results. The assessment time points varied between 6 and 16 weeks after randomisation. Considering that 50% to 70% of patients on antidepressants may use antidepressants for more than 2 years,[59 60] clinical trials with long-term follow-up are required to assess the full benefits and harms of duloxetine. This is particularly important for a medication class like antidepressants for which there is accumulating evidence of adverse effects of long-term use.[61–63] Another central area for improvement is the non-systematic assessment of adverse events in the included trials, which is expected

to lead to an underestimation of the prevalence of these effects.

Our review has other limitations. All trials included in this review were funded by an industry. Research has shown that industry involvement and funding introduce a systematic bias in favour of the medication which is not fully accounted for by the usual risk of bias assessments.[64] The bias arising from vested interests is often attributed to selective publication of positive results, underreporting of SAEs and harms as well as strict inclusion criteria for participants that might not reflect real-world settings.[65] Furthermore, all included trials were assessed to be at overall high risk of bias primarily due to the absence of published protocols, missing data and poor description of outcome assessment and blinding procedures. In addition, withdrawal effects are known to increase with increased duration of treatment, and the short periods of duloxetine treatment may underestimate the risk of a withdrawal syndrome for patients using these medications for longer periods, along with other adverse effects that only emerge after long-term use.[25] The certainty of evidence from all meta-analyses was low to very low. Hence, it is likely that our results overestimate the beneficial effects and underestimate the harmful effects of duloxetine.

CONCLUSION

Duloxetine appears to reduce depressive symptom scores and improve quality of life scores in the short term, but the effect sizes are minimal and of questionable patient importance. The short- and long-term effects of duloxetine on risks of SAEs and suicidality are uncertain. Duloxetine increases the risks of several short-term adverse events. Systematic assessments of benefits and harms over longer periods are required.

Author affiliations
[1]Copenhagen Trial Unit, Centre for Clinical Intervention Research, The Capital Region, Copenhagen University Hospital — Rigshospitalet, Copenhagen, Denmark
[2]Stolpegaard Psychotherapy Centre, Mental Health Services, The Capital Region, Gentofte, Denmark
[3]Department of Psychology, University of Copenhagen, Copenhagen, Denmark
[4]Department of Regional Health Research, The Faculty of Heath Sciences, University of Southern Denmark, Odense, Denmark
[5]Division of Psychiatry, University College London, London, UK
[6]Research and Development Department, North East London NHS Foundation Trust (NELFT), Essex, UK
[7](honorary position for MAH) Division of Psychiatry, University College London, London, UK
[8]Centre for Anaesthesiological Research, Department of Anaesthesiology, Zealand University Hospital, Køge, Denmark

X Joanna Moncrieff @joannamoncrieff and Mark Abie Horowitz @markhoro

Acknowledgements We thank Sarah Louise Klingenberg (Information Specialist, The Cochrane Hepato-Biliary Group, Copenhagen Trial Unit, Centre for Clinical Intervention Research, Copenhagen University Hospital – Rigshospitalet, Denmark) for the help with developing and conducting the search strategy.

Contributors CBK, SJ, FS, JM, MAH, KKK, MB, CG and JCJ contributed to the conceptualisation and design of the systematic review. FS and JJP screened studies for inclusion. FS, CBK, JJP, SJ and MB extracted data. FS and MM analysed data. FS and JCJ wrote the original draft. All authors commented on and approved the final manuscript. The corresponding author attests that all listed authors meet authorship criteria. JCJ is the guarantor of this research. JCJ affirms that this manuscript is an honest, accurate and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained. JCJ has access to all the data and takes responsibility for the integrity of the data and accuracy of the data analysis.

Funding The research was supported by The Copenhagen Trial Unit with salaries for the authors during their work on the review and writing of the manuscript.

Competing interests MAH and JM declare that they are collaborating investigators on the RELEASE trial – Redressing Long-term Antidepressant Use in General Practice – and the RELEASE+ trial in Australia funded by the National Health and Medical Research Council (NMHRC) and Medical Research Future Fund (MRFF). MAH reports that he is co-founder of Outro Health, a digital clinic helping people to safely stop unnecessary antidepressants in the US. JM is a co-investigator on a National Institute of Health Research (NIHR) funded study exploring methods of antidepressant discontinuation (REDUCE). She collects royalties from books on psychiatric drugs. All other authors have no known competing interests.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Authors are willing to share data on request and in accordance with data protection regulations and local information governance.

ORCID iDs
Faiza Siddiqui http://orcid.org/0000-0002-8358-6259
Johanne Juul Petersen http://orcid.org/0000-0001-9837-1958
Sophie Juul http://orcid.org/0000-0002-6171-2904
Caroline Barkholt Kamp http://orcid.org/0000-0002-7756-4694
Marija Barbateskovic http://orcid.org/0000-0001-8566-3660
Joanna Moncrieff http://orcid.org/0000-0003-1214-6974
Mark Abie Horowitz http://orcid.org/0000-0003-1318-2029
Christian Gluud http://orcid.org/0000-0002-8861-0799

REFERENCES

1 American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM–5), Available: https://dsm.psychiatryonline.org/dsmPreviousEditions
2 World Health Organization. Depression, Available: https://www.who.int/news-room/fact-sheets/detail/depression
3 Greenberg PE, Fournier AA, Sisitsky T, *et al*. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *J Clin Psychiatry* 2015;76:155–62.
4 IsHak WW, Mirocha J, James D, *et al*. Quality of life in major depressive disorder before/after multiple steps of treatment and one-year follow-up. *Acta Psychiatr Scand* 2015;131:51–60.
5 Saragoussi D, Christensen MC, Hammer-Helmich L, *et al*. Long-term follow-up on health-related quality of life in major depressive

disorder: a 2-year European cohort study. *Neuropsychiatr Dis Treat* 2018;14:1339–50.
6 Friedrich MJ. Depression Is the Leading Cause of Disability Around the World. *JAMA* 2017;317:1517.
7 Hillhouse TM, Porter JH. A brief history of the development of antidepressant drugs: from monoamines to glutamate. *Exp Clin Psychopharmacol* 2015;23:1–21.
8 Federal Drug Authority. Duloxetine (marketed as Cymbalta) Information, Available: https://www.fda.gov/drugs/postmarket-drug-safety-information-patients-and-providers/duloxetine-marketed-cymbalta-information
9 European Medical Association. Cymbalta-epar-product-information, Available: https://www.ema.europa.eu/en/documents/product-information/cymbalta-epar-product-information_en.pdf
10 Higgins JPT, Spiegelhalter DJ. Being sceptical about meta-analyses: a Bayesian perspective on magnesium trials in myocardial infarction. *Int J Epidemiol* 2002;31:96–104.
11 Jakobsen JC, Katakam KK, Schou A, *et al*. Selective serotonin reuptake inhibitors versus placebo in patients with major depressive disorder. A systematic review with meta-analysis and Trial Sequential Analysis. *BMC Psychiatry* 2017;17:58.
12 Cipriani A, Furukawa TA, Salanti G, *et al*. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet* 2018;391:1357–66.
13 Hengartner MP, Plöderl M. Newer-Generation Antidepressants and Suicide Risk in Randomized Controlled Trials: A Re-Analysis of the FDA Database. *Psychother Psychosom* 2019;88:247–8.
14 Siddiqui F, Barbateskovic M, Juul S, *et al*. Duloxetine versus 'active' placebo, placebo or no intervention for major depressive disorder; a protocol for a systematic review of randomised clinical trials with meta-analysis and trial sequential analysis. *Syst Rev* 2021;10:171.
15 PRISMA 2020, 2024.
16 Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;23:56–62.
17 National Institute for Clinical Excellence. Depression: Management of Depression in Primary and Secondary Care. Clinical Practice Guideline no 23. 2004.
18 Hengartner MP, Plöderl M. Estimates of the minimal important difference to evaluate the clinical significance of antidepressants in the acute treatment of moderate-to-severe depression. *BMJ Evid Based Med* 2022;27:69–73.
19 International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Guideline: Integrated Addemdum to ICH E6(R1): Guideline for Good Clinical Practice (ICH-GCP), Available: https://ichgcp.net/
20 Sterne JAC, Savović J, Page MJ, *et al*. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.
21 IntHout J, Ioannidis JPA, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med Res Methodol* 2014;14:25.
22 RStudio. RStudio: Integrated Development for R. 2020. Available: http://www.rstudio.com
23 Higgins JPT, Thomas J, Chandler J, *et al*. *Cochrane handbook for systematic reviews of interventions, 2nd edn (Cochrane book series)*. John Wiley & Sons, 2019.
24 Jakobsen JC, Wetterslev J, Winkel P, *et al*. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Med Res Methodol* 2014;14:120.
25 Horowitz MA, Framer A, Hengartner MP, *et al*. Estimating Risk of Antidepressant Withdrawal from a Review of Published Data. *CNS Drugs* 2023;37:143–57.
26 Wetterslev J, Thorlund K, Brok J, *et al*. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *J Clin Epidemiol* 2008;61:64–75.
27 Thorlund K, Engstrøm J, Wetterslev J, *et al*. User manual for Trial Sequential Analysis (TSA), Available: www.ctu.dk/tsa
28 Brok J, Thorlund K, Gluud C, *et al*. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. *J Clin Epidemiol* 2008;61:763–9.
29 Baldwin DS, Loft H, Dragheim M. A randomised, double-blind, placebo controlled, duloxetine-referenced, fixed-dose study of three dosages of Lu AA21004 in acute treatment of major depressive disorder (MDD). *Eur Neuropsychopharmacol* 2012;22:482–91.
30 Boulenger JP, Loft H, Olsen CK. Efficacy and safety of vortioxetine (Lu AA21004), 15 and 20 mg/day: a randomized, double-blind, placebo-controlled, duloxetine-referenced study in the acute treatment of adult patients with major depressive disorder. *Int Clin Psychopharmacol* 2014;29:138–49.
31 Cutler AJ, Montgomery SA, Feifel D, *et al*. Extended release quetiapine fumarate monotherapy in major depressive disorder: a placebo- and duloxetine-controlled study. *J Clin Psychiatry* 2009;70:526–39.
32 Detke MJ, Lu Y, Goldstein DJ, *et al*. Duloxetine, 60 mg once daily, for major depressive disorder: a randomized double-blind placebo-controlled trial. *J Clin Psychiatry* 2002;63:308–15.
33 Detke MJ, Lu Y, Goldstein DJ, *et al*. Duloxetine 60 mg once daily dosing versus placebo in the acute treatment of major depression. *J Psychiatr Res* 2002;36:383–90.
34 Detke MJ, Wiltse CG, Mallinckrodt CH, *et al*. Duloxetine in the acute and long-term treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. *Eur Neuropsychopharmacol* 2004;14:457–70.
35 Goldstein DJ, Lu Y, Detke MJ, *et al*. Duloxetine in the treatment of depression: a double-blind placebo-controlled comparison with paroxetine. *J Clin Psychopharmacol* 2004;24:389–99.
36 Goldstein DJ, Mallinckrodt C, Lu Y, *et al*. Duloxetine in the treatment of major depressive disorder: a double-blind clinical trial. *J Clin Psychiatry* 2002;63:225–31.
37 Katona C, Hansen T, Olsen CK. A randomized, double-blind, placebo-controlled, duloxetine-referenced, fixed-dose study comparing the efficacy and safety of Lu AA21004 in elderly patients with major depressive disorder. *Int Clin Psychopharmacol* 2012;27:215–23.
38 Mahableshwarkar AR, Jacobsen PL, Chen Y. A randomized, double-blind trial of 2.5 mg and 5 mg vortioxetine (Lu AA21004) versus placebo for 8 weeks in adults with major depressive disorder. *Curr Med Res Opin* 2013;29:217–26.
39 Mahableshwarkar AR, Jacobsen PL, Chen Y, *et al*. A randomized, double-blind, duloxetine-referenced study comparing efficacy and tolerability of 2 fixed doses of vortioxetine in the acute treatment of adults with MDD. *Psychopharmacology (Berl)* 2015;232:2061–70.
40 Mahableshwarkar AR, Zajecka J, Jacobson W, *et al*. A Randomized, Placebo-Controlled, Active-Reference, Double-Blind, Flexible-Dose Study of the Efficacy of Vortioxetine on Cognitive Function in Major Depressive Disorder. *Neuropsychopharmacology* 2015;40:2025–37.
41 Nierenberg AA, Greist JH, Mallinckrodt CH, *et al*. Duloxetine versus escitalopram and placebo in the treatment of patients with major depressive disorder: onset of antidepressant action, a non-inferiority study. *Curr Med Res Opin* 2007;23:401–16.
42 Oakes TMM, Myers AL, Marangell LB, *et al*. Assessment of depressive symptoms and functional outcomes in patients with major depressive disorder treated with duloxetine versus placebo: primary outcomes from two trials conducted under the same protocol. *Hum Psychopharmacol* 2012;27:47–56.
43 Perahia DGS, Wang F, Mallinckrodt CH, *et al*. Duloxetine in the treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. *Eur Psychiatry* 2006;21:367–78.
44 Raskin J, Wiltse CG, Siegal A, *et al*. Efficacy of duloxetine on cognition, depression, and pain in elderly patients with major depressive disorder: an 8-week, double-blind, placebo-controlled trial. *Am J Psychiatry* 2007;164:900–9.
45 Robinson M, Oakes TM, Raskin J, *et al*. Acute and long-term treatment of late-life major depressive disorder: duloxetine versus placebo. *Am J Geriatr Psychiatry* 2014;22:34–45.
46 Tourian KA, Padmanabhan SK, Groark J, *et al*. Desvenlafaxine 50 and 100 mg/d in the treatment of major depressive disorder: an 8-week, phase III, multicenter, randomized, double-blind, placebo-controlled, parallel-group trial and a post hoc pooled analysis of three studies. *Clin Ther* 2009;31 Pt 1:1405–23.
47 NIH. 6-week study treatment to evaluate the safety and effectiveness of AZD2066 in patients with major depressive disorder, NCT01145755. 2012.
48 Kamijima K. Clinical evaluation of duloxetine in the treatment of major depressive disorder placebo and paroxetine controlled double-blind comparative study. *Japanese J Clin Psychopharmacol* 2009;12:1613–34.
49 Eli Lilly and Company. Randomised, double-blind, parallel-group, placebo-controlled, duloxetine-referenced dose-finding study of Lu AA24530 in Major Depressive Disorder (11918A). Synopsis. 2010.
50 Eli Lilly and Company. Validation of Daily Telephone Self-Assessment in the Study of Antidepressant Treatment Outcome (H8I-MC-HQAC). Clinical Study Summary. 2007.
51 Eli Lilly and Company. A double-blind, placebo- and clomipramine-controlled study of duloxetine in patients with major depression (F1J-MC-HMAI). Abbreviated clinical study report. 2001.
52 Eli Lilly and Company. Duloxetine versus Placebo in the Treatment of Major Depression (F1J-MC-HMAQ(B)). Clinical Study Report. 2001.

53 Eli Lilly and Company. Duloxetine Versus Placebo and Paroxetine in the Acute Treatment of Major Depression (F1J-MC-HMAT), Clinical Study Report. 2001.
54 Eli Lilly and Company. Duloxetine/placebo in Major Depression (F1J-MC-HMAG). 1996.
55 Eli Lilly and Company. Duloxetine 20/30 mg vs. Placebo in Major Depression (F1J-MC-HMAH). Clinical Study Report. 1996.
56 Leucht S, Fennema H, Engel R, *et al*. What does the HAMD mean? *J Affect Disord* 2013;148:243–8.
57 Bagby RM, Ryder AG, Schuller DR, *et al*. The Hamilton Depression Rating Scale: has the gold standard become a lead weight? *Am J Psychiatry* 2004;161:2163–77.
58 Paludan-Müller AS, Sharma T, Rasmussen K, *et al*. Extensive selective reporting of quality of life in clinical study reports and publications of placebo-controlled trials of antidepressants. *Int J Risk Saf Med* 2021;32:87–99.
59 Mojtabai R, Olfson M. National Trends in Long-Term Use of Antidepressant Medications. *J Clin Psychiatry* 2014;75:169–77.
60 Johnson CF, Macdonald HJ, Atkinson P, *et al*. Reviewing long-term antidepressants can reduce drug burden: a prospective observational cohort study. *Br J Gen Pract* 2012;62:e773–9.
61 Kinrys G, Gold AK, Pisano VD, *et al*. Tachyphylaxis in major depressive disorder: A review of the current state of research. *J Affect Disord* 2019;245:488–97.
62 Laporte S, Chapelle C, Caillet P, *et al*. Bleeding risk under selective serotonin reuptake inhibitor (SSRI) antidepressants: A meta-analysis of observational studies. *Pharmacol Res* 2017;118:19–32.
63 Moncrieff J. Persistent adverse effects of antidepressants. *Epidemiol Psychiatr Sci* 2020;29:e56.
64 Lundh A, Lexchin J, Mintzes B, *et al*. Industry sponsorship and research outcome. *Cochrane Database Syst Rev* 2017;2.
65 Hengartner MP. Methodological Flaws, Conflicts of Interest, and Scientific Fallacies: Implications for the Evaluation of Antidepressants' Efficacy and Harm. *Front Psychiatry* 2017;8:275.