RESEARCH ARTICLE

Open Access

Selective serotonin reuptake inhibitors versus placebo in patients with major depressive disorder. A systematic review with meta-analysis and Trial Sequential Analysis

Janus Christian Jakobsen[1,2*], Kiran Kumar Katakam[1], Anne Schou[1], Signe Gade Hellmuth[1], Sandra Elkjær Stallknecht[1], Katja Leth-Møller[1], Maria Iversen[1], Marianne Bjørnø Banke[1], Iggiannguaq Juhl Petersen[1], Sarah Louise Klingenberg[1], Jesper Krogh[3], Sebastian Elgaard Ebert[1], Anne Timm[1], Jane Lindschou[1] and Christian Gluud[1]

Abstract

Background: The evidence on selective serotonin reuptake inhibitors (SSRIs) for major depressive disorder is unclear.

Methods: Our objective was to conduct a systematic review assessing the effects of SSRIs versus placebo, 'active' placebo, or no intervention in adult participants with major depressive disorder. We searched for eligible randomised clinical trials in The Cochrane Library's CENTRAL, PubMed, EMBASE, PsycLIT, PsycINFO, Science Citation Index Expanded, clinical trial registers of Europe and USA, websites of pharmaceutical companies, the U.S. Food and Drug Administration (FDA), and the European Medicines Agency until January 2016. All data were extracted by at least two independent investigators. We used Cochrane systematic review methodology, Trial Sequential Analysis, and calculation of Bayes factor. An eight-step procedure was followed to assess if thresholds for statistical and clinical significance were crossed. Primary outcomes were reduction of depressive symptoms, remission, and adverse events. Secondary outcomes were suicides, suicide attempts, suicide ideation, and quality of life.

(Continued on next page)

* Correspondence: jcj@ctu.dk
[1]The Copenhagen Trial Unit, Centre for Clinical Intervention Research, Department 7812 Rigshospitalet, Copenhagen University Hospital, Blegdamsvej 9, Rigshospitalet, DK 2100 Copenhagen, Denmark
[2]Department of Cardiology, Holbæk Hospital, Holbæk, Denmark
Full list of author information is available at the end of the article

(Continued from previous page)

Results: A total of 131 randomised placebo-controlled trials enrolling a total of 27,422 participants were included. None of the trials used 'active' placebo or no intervention as control intervention. All trials had high risk of bias. SSRIs significantly reduced the Hamilton Depression Rating Scale (HDRS) at end of treatment (mean difference −1.94 HDRS points; 95% CI −2.50 to −1.37; $P < 0.00001$; 49 trials; Trial Sequential Analysis-adjusted CI −2.70 to −1.18); Bayes factor below predefined threshold ($2.01*10^{-23}$). The effect estimate, however, was below our predefined threshold for clinical significance of 3 HDRS points. SSRIs significantly decreased the risk of no remission (RR 0.88; 95% CI 0.84 to 0.91; $P < 0.00001$; 34 trials; Trial Sequential Analysis adjusted CI 0.83 to 0.92); Bayes factor (1426.81) did not confirm the effect). SSRIs significantly increased the risks of serious adverse events (OR 1.37; 95% CI 1.08 to 1.75; $P = 0.009$; 44 trials; Trial Sequential Analysis-adjusted CI 1.03 to 1.89). This corresponds to 31/1000 SSRI participants will experience a serious adverse event compared with 22/1000 control participants. SSRIs also significantly increased the number of non-serious adverse events. There were almost no data on suicidal behaviour, quality of life, and long-term effects.

Conclusions: SSRIs might have statistically significant effects on depressive symptoms, but all trials were at high risk of bias and the clinical significance seems questionable. SSRIs significantly increase the risk of both serious and non-serious adverse events. The potential small beneficial effects seem to be outweighed by harmful effects.

Systematic review registration: PROSPERO CRD42013004420.

Keywords: Depression, SSRI, Systematic review

Background

Selective serotonin reuptake inhibitors (SSRIs) are often first-line treatment for depression and prescriptions for SSRIs are increasing [1, 2]. A number of reviews with meta-analysis have assessed the effects of SSRIs in adults with major depressive disorder [3–8], generally concluding that SSRIs have a statistically significant effect on depressive symptoms [3–8]. However, the results of the reviews have been limited by not using predefined Cochrane methodology [3–8], only including subgroups of depressed patients [9, 10], not searching all relevant databases [3–8, 10], not systematically assessing harms [3–8, 10], and not systematically assessing risks of bias [3–8, 10]. We have summarised the characteristics and the results of previous systematic reviews in Table 1. Accordingly, the evidence on the effects of SSRIs is unclear. Using, e.g., a composite outcome of all serious adverse events (according to ICH-GCP [11]) might show how SSRIs work. Furthermore, assessments of quality of life might demonstrate if SSRIs have clinically meaningful effects. It is of utmost importance to assess the clinical significance of review results if statistically significant results are shown [12, 13].

Our objective was to conduct a comprehensive systematic review assessing the beneficial and harmful effects of SSRIs versus placebo, 'active' placebo, or no intervention in adult participants with major depressive disorder using our eight-step procedure for assessing evidence in systematic reviews [13].

Methods

Details regarding the methodology are described in our protocol, which was registered prior to the systematic literature searches [14]. The methodology was not changed after the analysis of the review results began [14].

We included all randomised clinical trials comparing the effects of SSRIs (citalopram; escitalopram; sertraline; fluoxetine; paroxetine; or fluvoxamine) versus placebo, 'active placebo' (any active substance employed to mimic the adverse effects of taking a SSRI) [15], or no intervention. We also planned to perform subgroup analyses comparing the effects of the different doses (see Subgroup analyses). If a trial had three arms (e.g., a three-arm trial randomising the participants to two different SSRIs and placebo) then we divided the total number of control participants with two but kept the means and SDs in each group unchanged [16, 17].

Independent investigators searched for eligible trials published before January 2016 in The Cochrane Library's CENTRAL, PubMed, EMBASE, PsychLIT, PsycINFO, clinicaltrials.gov., and Science Citation Index Expanded [14] (see Additional file 1: Search strategies). Trials were included irrespective of language, publication status, publication year, and publication type. To identify unpublished trials, we searched clinical trial registers of Europe and USA, websites of pharmaceutical companies, websites of U.S. Food and Drug Administration (FDA) and European Medicines Agency, and we requested the U.S. Food and Drug Administration (FDA) to provide all publicly releasable information about relevant clinical trials of SSRIs that were submitted for marketing approval.

Participants had to be 18 years or older and have a primary diagnosis of major depressive disorder based on standardised criteria, such as DSM III, DSM III-R, DSM IV, DSM V, or ICD 10 [14].

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 3 of 28

Table 1 Overview of previous reviews

	First author	Title	Year of publication	Design	Type of SSRI assessed	Information sources	No. of trials	No. of patients	Published protocol	Assessment of adverse events	Assessment of risk of bias
Reviews concluding that SSRIs have beneficial effect on major depressive disorder	Gibbons et al.	Benefits From Antidepressants: Synthesis of 6-Week Patient-Level Outcomes From Double-blind Placebo-Controlled Randomized Trials of Fluoxetine and Venlafaxine	2012	Patient level meta-analysis	Fluoxetine	Eli Lilly and Co	16	3595	No	No	No
	Undurraga et al.	Randomized, Placebo-Controlled Trials of Antidepressants for Acute Major Depression: Thirty-Year Meta-Analytic Review	2011	Systematic review	Fluoxetine, sertraline, paroxetine, citalopram, escitalopram	Medline, CINAH Library, Cochrane Library, PsycINFO	51	5285	No	No	Only publication bias
	Wilson et al.	Antidepressants Versus Placebo for the Depressed Elderly	2001	Cochrane review	Fluoxetine	PsycLIT, MEDLINE, EMBASE, LILACS, CINAHL, SIGLE, Psyndex, National Research Register, Dissertation Abstracts International	2	365	Yes	No	Only allocation concealment
	Aroll et al.	Antidepressants versus placebo for depression in primary care (Review)	2009	Cochrane review	Sertraline, escitalopram, citalopram	CCDANCTR	4	707	Yes	Yes	Only allocation concealment (QRS)
Reviews concluding that SSRIs have no effect on mild to moderate depression but have beneficial effect on severe depression	Fournier et al.	Antidepressant Drug Effects and Depression Severity: A Patient-Level Meta-analysis	2010	Patient level meta-analysis	Paroxetine	PubMed, PsycINFO, Cochrane Library	3	240	No	No	No
	Khan et al.	Severity of Depression and Response to Antidepressants and Placebo: An Analysis of the Food and Drug Administration Database	2002	Systematic review	Fluoxetine, sertraline, paroxetine	FDA	18	Unclear	No	No	No
Reviews concluding that SSRIs have questionable effect on major depressive disorder	Turner et al.	Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy	2008	Systematic review	Fluoxetine, sertraline, paroxetine, citalopram, escitalopram	FDA, PubMed, Cochrane Library	38	Unclear	No	No	Only publication bias
	Kirsch et al.	Initial Severity and Antidepressant Benefits: A Meta-Analysis of Data Submitted to the Food and Drug Administration	2008	Systematic review	Fluoxetine, sertraline, paroxetine, citalopram	FDA, PubMed	21	1708	No	No	Only publication bias
Reviews not showing anything about the effects of SSRIs on major depressive disorder	Moncrieff et al.	Active Placebos Versus Antidepressants for Depression (Review)	2012	Cochrane review	No SSRIs studied, only TCAs	CCDANCTR			Yes		

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 4 of 28

We excluded trials specifically randomising depressed participants with: a specific somatic disease, schizophrenia, or depression during or after pregnancy.

Outcomes
Primary outcomes

- Depressive symptoms measured on the 17-item or 21-item Hamilton Depression Rating Scale (HDRS) [18], the Montgomery-Asberg Depression Rating Scale (MADRS) [19], or the Beck's Depression Inventory (BDI) [20].
- Remission (Hamilton <8 points; BDI <10 points; MADRS <10 points).
- Adverse events during the intervention period which were classified as serious and non-serious adverse events [11]. Serious adverse events were defined as medical events that were life threatening, resulted in death, disability, or significant loss of function, or caused hospital admission or prolonged hospitalisation [11]. The remaining events were classified as non-serious adverse events [11].

Secondary outcomes

- Suicides, suicide attempts, and suicide ideation during the intervention period.
- Quality of life (scale used by the trialists).

The time point of primary interest was end of treatment (defined by trialist) [14]. We also planned to report results assessed at maximum follow-up [14].

Handling of missing data
If the participant was not included in the analysis of 'no remission', we assumed that they had 'no remission' [17]. If, e.g. 23/50 participants had 'no remission' but a total of 53 were randomised then we assumed that 26 had 'no remission'. For all remaining outcomes we used observed data when these were reported [17].

Subgroup analyses
We planned the following subgroup analyses: participants with HDRS baseline ≤23 compared to >23 points; different SSRIs; trials using a placebo washout period before randomisation compared to trials using no washout period; participants with alcohol or drug dependence compared to no dependence; elderly participants compared to younger participants; duration of treatment below 8 weeks compared to equal to or above 8 weeks; and SSRI below or equal to median dose compared to above median dose.

We planned to assess if SSRIs differ according to psychiatric comorbidities [17]. Four trials included depressed participants with comorbid anxiety but none of these trials reported HDRS so this could not be performed. No other comorbidities (including borderline personality disorder, chronic depression, and treatment resistant depression) were identified in the included trials. We planned to assess if the effects of SSRIs differed: (1) when the SSRI was delivered as add-on therapy to another antidepressant drug; (2) per use of different forms of control interventions ('active placebo', traditional placebo, and no intervention); and (3) if electroconvulsive therapy was used as co-intervention. However, these analyses were not possible.

We used test for subgroup differences to assess if the effects of SSRIs seemed to differ between the different types of participants if either a trial specifically randomised a certain type of participants, or a trial reported results separately for each specific type of participants.

Selection of trials and data extraction
Review authors (KKK, AS, SGH, SES, KLM, MI, MBB, IJP, JK, SLK, AT, SE, JCJ) worked in pairs and independently selected relevant trials and extracted data. A standardised data extraction sheet was used (see Data extraction form). If a trial was identified by only one, it was discussed whether the trial should be included. In case of discrepancy, a third review author (JCJ) was consulted. We contacted review authors if relevant information was missing.

Assessment of the statistical and clinical significance
Our methodology was based on The Cochrane Handbook and GRADE [14, 16, 21]. We assessed statistical and clinical significance according to our eight-step procedure [13]:

(1) We obtained 95% confidence intervals (95% CI) and *P*-values from all planned random-effects [22] and fixed-effect meta-analyses [16] and reported the most conservative result as the main result [13]. Review Manager version 5.3 was used for all meta-analyses [23].
(2) Sensitivity analyses and subgroup analyses were conducted to explore the reasons for substantial statistical heterogeneity [13, 16]. Statistical heterogeneity was assessed by visual inspection of forest plots and by the heterogeneity (I^2 or D^2) [13, 16, 24, 25].
(3) We defined three primary outcomes in our protocol [14]. Our threshold for significance was therefore adjusted according to problems with multiplicity [13] by dividing 0.05 with the value halfway between 1 (no adjustment) and 3 (Bonferroni adjustment) [13, 17] resulting in 0.05/2 = 0.025.
(4) Cumulative meta-analyses are at risk of producing random errors due to sparse data and multiple

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 5 of 28

testing of accumulating data [25, 26]. Therefore, Trial Sequential Analysis version 0.9.5.5 beta was applied to control this risk (http://www.ctu.dk/tsa/) [27]. The required information size (that is the number of participants needed in a meta-analysis to detect or reject a certain intervention effect) was calculated [25, 28]. The required information size is based upon the event proportion in the control group; the assumption of a plausible relative risk (RR) reduction; and the assumed heterogeneity or diversity of the meta-analysis [25, 29]. Trial Sequential Analysis enables testing to be conducted each time a new trial is included in the meta-analysis. Based on the required information size, trial sequential monitoring boundaries are constructed. This enables one to determine the statistical inference concerning cumulative meta-analysis that has not yet reached the required information size [25]. Firm evidence may be established if one of the trial sequential monitoring boundaries (for benefit (upper red dotted line), for harm (lower red dotted line), or for futility (vertical red dotted line)) is crossed by the cumulative Z-score before reaching the required information size, in which case further trials may turn out to be superfluous [28]. Trial Sequential Analysis- adjusted confidence intervals are also presented [28]. For dichotomous outcomes, we estimated the diversity-adjusted required information size based on the proportion of patients with an outcome in the control group, a relative risk reduction of 30%, an alpha of 2.5%, a beta of 20%, and diversity in the trials [13, 17]. For continuous outcomes, we estimated the required information size based on a HDRS mean difference of three points, the standard deviation observed in the control group, an alpha of 2.5%, a beta of 20%, and diversity in the trials [13, 17]. All outcomes were assessed with Trial Sequential Analysis [13, 24, 25].

(5) We calculated Bayes factors for all primary outcomes. A low *P* value indicates that an observed result is unlikely given that there is no difference in effect between the compared intervention groups (i.e., the null hypothesis is true) [12, 13]. Even very low *P* values may, therefore, be misleading because the probability that the actual measured difference in effect of the compared interventions resulted from an a priori anticipated 'true' difference needs to be considered [13]. For this purpose, it is helpful to calculate Bayes factor for the primary outcomes. It will show the ratios between the *P*-value probabilities of the meta-analysis result divided by the probability of the meta-analysis result given that an anticipated intervention effect is the true effect [12, 13]. In other words, the lower the Bayes factor

gets the more confident one should be that an actual intervention effect (the anticipated intervention effect) has produced the meta-analysis results and that a given difference between the compared groups is not caused by random error [12, 13]. Calculation of Bayes factor is not part of standard Cochrane methodology.

(6) We assessed the potential impact of bias on the review results [16]. To assess the potential impact of missing data (incomplete outcome data bias) we assessed a 'best-worst' case scenario assuming that all participants lost to follow-up in the SSRI group had a beneficial outcome (the group mean plus 1 standard deviation (SD) or plus 2 SDs); and all those with missing outcomes in the placebo group have had a harmful outcome (the group mean minus 1 SD or minus 2 SDs) [13, 17]. We also performed the reverse 'worst-best-case' scenario analysis [13, 17].

(7) We assessed the risk of publication bias by visual inspection of funnel plots and by tests for funnel plot asymmetry [13, 16, 30].

(8) We assessed clinical significance of our results. As previously suggested [4, 8, 31], we chose a drug-placebo difference of 3 points on the 17-item HDRS or an effect size of 0.50 SMD as the threshold for clinical significance (see Discussion) [14].

Results

We have summarised the selection of trials in Fig. 1 and excluded trials in Additional file 2: List of excluded trials.

Using our strict inclusion and exclusion criteria, a total of 195 publications/unpublished trials were identified and included. Due to multiple publications of single trials and lack of useful data, only 131 trials randomised clinical trials [32–164] enrolling a total of 27,422 participants were included in our analyses. 116 were published trials [32–118, 120–131, 133–136, 141, 144, 146–148, 152, 153, 156–164] and 15 were unpublished trials [102, 119, 132, 137–140, 142, 143, 145, 149–151, 155, 165]. Of the 15 unpublished trials, eight were identified via company websites, two via clinicaltrials.gov, and four via FDA (see Additional file 3: Trial Characteristics).

We did not identify any trials using 'active placebo' or 'no intervention' as control interventions. Most trials used broad inclusion criteria and randomised adult depressed women and men below 65 years and used SSRI for 6 to 12 weeks. Twelve of the included trials specifically randomised elderly (most often participants >65 years) depressed participants [45, 48, 54, 63, 82, 110, 111, 113, 122, 144, 162, 166], five trials randomised depressed alcohol dependant participants [47, 61, 116, 120, 156], and three trials randomised depressed opioid and cocaine dependant participants [44, 115, 164].

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 6 of 28

Fig. 1 PRISMA flowchart
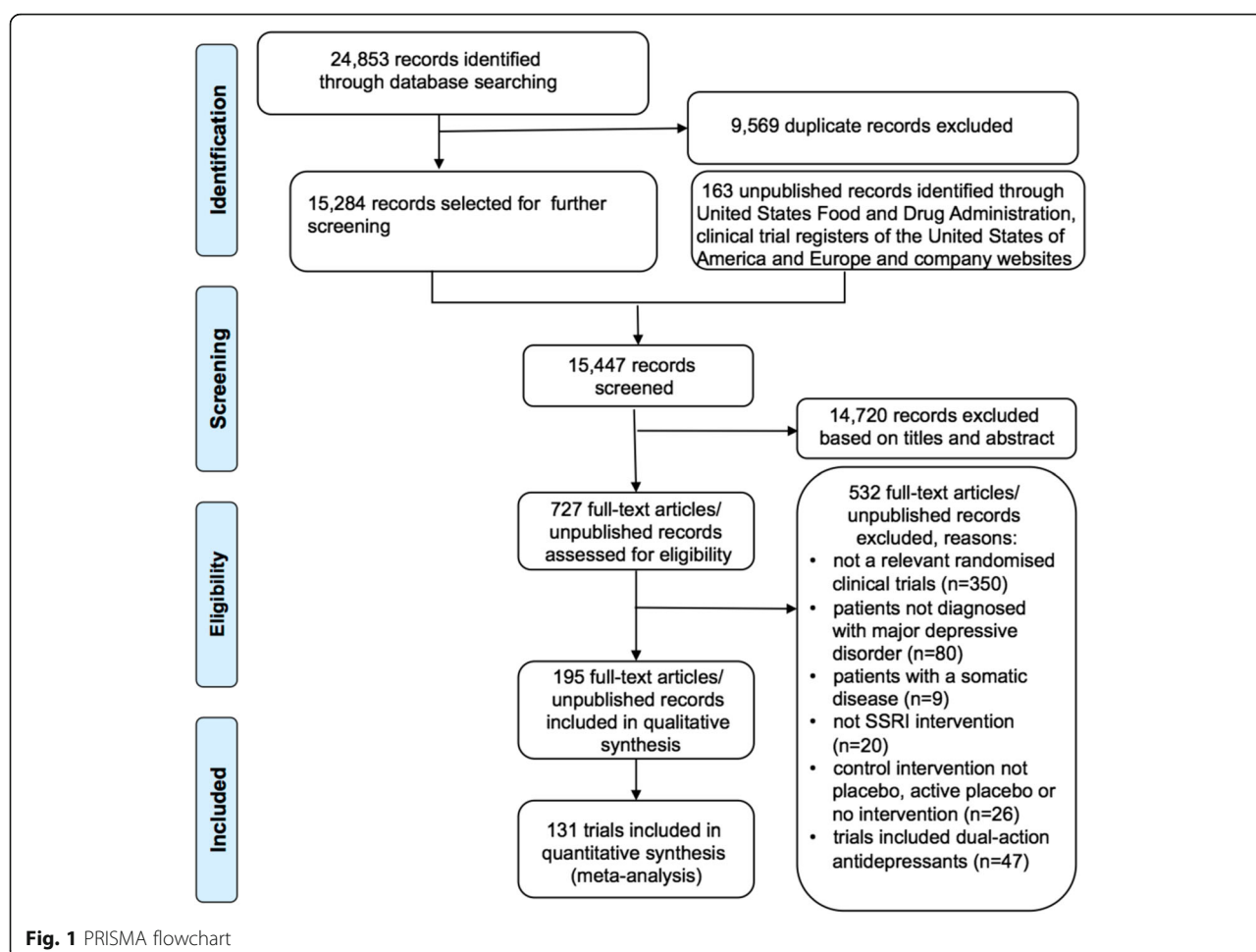
Types of the randomised participants and other trial characteristics are summarised in Additional file 3: Trial characteristics.

Primary outcomes

Hamilton depression rating scale (HDRS)

Twenty-two trials reported mean HDRS scores and standard deviation (SD) [32–50, 103, 144, 162] and 27 trials reported mean HDRS change scores and SD at end of treatment [51, 52, 54–66, 121, 123, 137, 138, 141, 142, 145, 146, 167, 168]. Random-effects meta-analysis of these 49 trials showed that SSRIs versus placebo significantly reduced the HDRS score (mean difference −1.94 points; 95% CI −2.50 to −1.37; $P < 0.00001$) (Fig. 2). Twenty-four trials reported only mean HDRS scores or presented a graph showing the mean HDRS scores, but did not report the SD at end of treatment [67–88, 166]. We planned to impute missing SDs based on observed standard deviations from trials with similar characteristics [17]. Trial characteristics, sample sizes, and statistical weight of the included trials were similar across trials and we therefore chose to impute the missing SDs

with a value of 8 points (the mean of the observed standard deviations rounded up to the nearest integer). Nineteen trials reported only mean HDRS change scores or presented a graph showing the mean change HDRS scores, but did not report the SD [89–102, 108, 143, 157, 158, 169]. We imputed the missing SDs with a value of 7 points (the mean of the observed standard deviations rounded up to the nearest integer) [17]. Random-effects meta-analysis of the results of all 92 trials showed that SSRIs versus placebo significantly reduced the HDRS score (mean difference −2.25 points; 95% CI −2.69 to −1.83; $P < 0.00001$).

Heterogeneity

The primary meta-analysis showed statistically significant heterogeneity ($I^2 = 75\%$; $P < 0.00001$) (Fig. 2). Visual inspection of the forest plot indicated that one trial seemed to have a more extreme effect (larger intervention effect estimate and smaller confidence interval) [58]. For exploratory purposes, we tried to exclude this trial from the analysis and this reduced the I^2 to 29%, but removing this trial did not substantially alter the meta-analysis result

Fig. 2 Random-effects meta-analysis of the results on the Hamilton depression rating scale
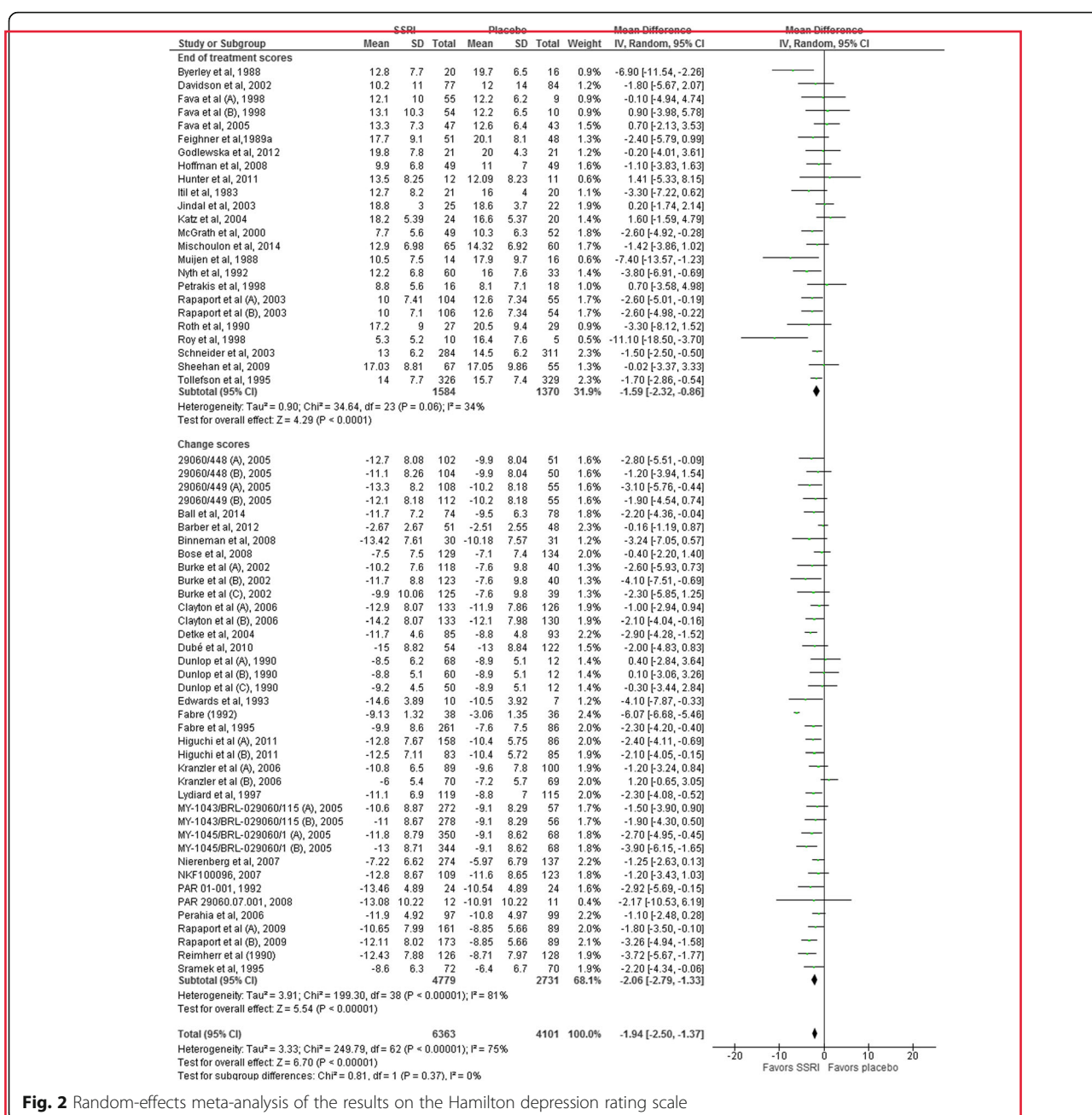
(mean difference after removing the trial from the analysis −1.77 HDRS points; 95% CI −2.12 to −1.42).

Subgroup analysis per risk of bias

All the included trials had high risk of bias. Hence, it was not possible to perform a subgroup analysis of trials at low risk of bias (Fig. 3) [13, 17]. However, four trials [33, 60, 103, 121] were classified as *potentially lower* risk of bias (based on bias risk assessment of 'generating allocation sequence', 'allocation concealment', 'blinding of participants and treatment providers', and 'blinding of

outcome assessment') [13, 17]. Random-effects meta-analysis of these four trials showed that SSRIs versus placebo reduced the HDRS score (mean difference −2.07 points; 95% CI −3.06 to −1.08). Test for subgroup differences between trials with *lower* risk of bias ($n = 4$) compared to trials with high risk of bias ($n = 45$) was not significant ($P = 0.82$).

Incomplete outcome data

Meta-analysis of the best-worst case scenario analyses adding 1 SD (mean difference −3.38 HDRS points in

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 8 of 28

Fig. 3 Risk of bias in the included randomized clinical trials
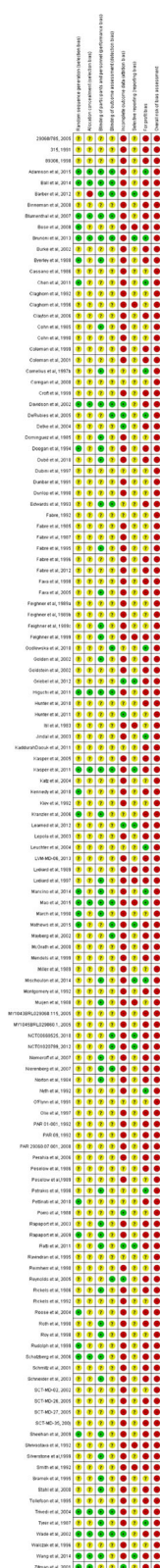
favour of SSRI; 95% CI −4.10 to −2.66) and adding 2 SDs (mean difference −4.50 HDRS points in favour of SSRI; 95% CI −5.37 to −3.63) for missing values showed large significant intervention effect estimates [13, 17]. Meta-analysis of the worst-best case scenario analyses showed a significant intervention effect estimate when adding 1 SD for missing values (mean difference −0.77 points; 95% CI −1.45 to −0.09) and no significant intervention effects when adding 2 SDs for missing values (mean difference 0.46 points; 95% CI −0.38 to 1.30).

Other subgroup analyses

Meta-analysis of the results of the 26 trials with a mean baseline HDRS score >23 points showed a mean difference of −2.69 HDRS points; 95% CI −3.59 to −1.78; $P < 0.00001$ [32, 35–38, 41, 43, 45–47, 49, 52, 57–59, 63–65, 121, 137, 142, 145, 167, 170, 171]. Meta-analysis of the results of the 20 trials with a mean baseline HDRS score ≤23 points showed a mean difference of -1.29 HDRS points; 95% CI −1.76 to −0.82; $P < 0.00001$ [33, 34, 39, 40, 44, 48, 50, 51, 54, 55, 60–63, 103, 141, 144, 146, 162, 168]. Test for subgroup difference was significant ($I^2 =$ 86.2%; $P = 0.007$). We performed post-hoc meta-regression (STATA 14) with baseline HDRS as a covariate in the meta-analysis. This analysis showed that the effects of SSRIs seem to increase with larger baseline HDRS scores (coefficient −0.33 points; 95% CI −0.44 to −0.22; $P < 0.0001$).

The following tests for subgroup differences did not show any significant differences: trials assessing the effects of the different SSRIs (number of trials 49; $I^2 =$ 2.2%; $P = 0.40$) (Fig. 4); published trials (45 trials) compared to unpublished trials (4 trials): $I^2 = 25.2\%$; $P = 0.25$; trials randomising elderly participants (6 trials) compared to younger participants (43 trials): $I^2 = 0\%$; $P = 0.94$ (Fig. 5); trials with washout period (40 trials) compared to trials without washout period (9 trials): $I^2 = 63.6\%$; $P = 0.10$ (Fig. 6); trials randomising drug or alcohol dependant participants (3 trials) compared to the remaining trials (46 trials): $I^2 = 0\%$; $P = 0.58$; trials with an intervention period below 8 weeks (19 trials) compared to the remaining trials (30 trials): $I^2 = 36.6$; $P = 0.21$; and dose of the chosen SSRI (dose below the median (6 trials) compared to equal to or above the median (9 trials) of the SSRI: $I^2 = 0\%$; $P = 0.65$. The latter two subgroup analyses were post-hoc analyses.

We performed an additional post-hoc subgroup analysis comparing trials with low risk of financial bias to trials with high risk of financial bias (Additional file 4: Figure S1). Test for subgroup differences showed no significant difference ($P = 0.18$). When the four trials with low risk of bias of financial bias were analysed separately then there was no significant difference between the
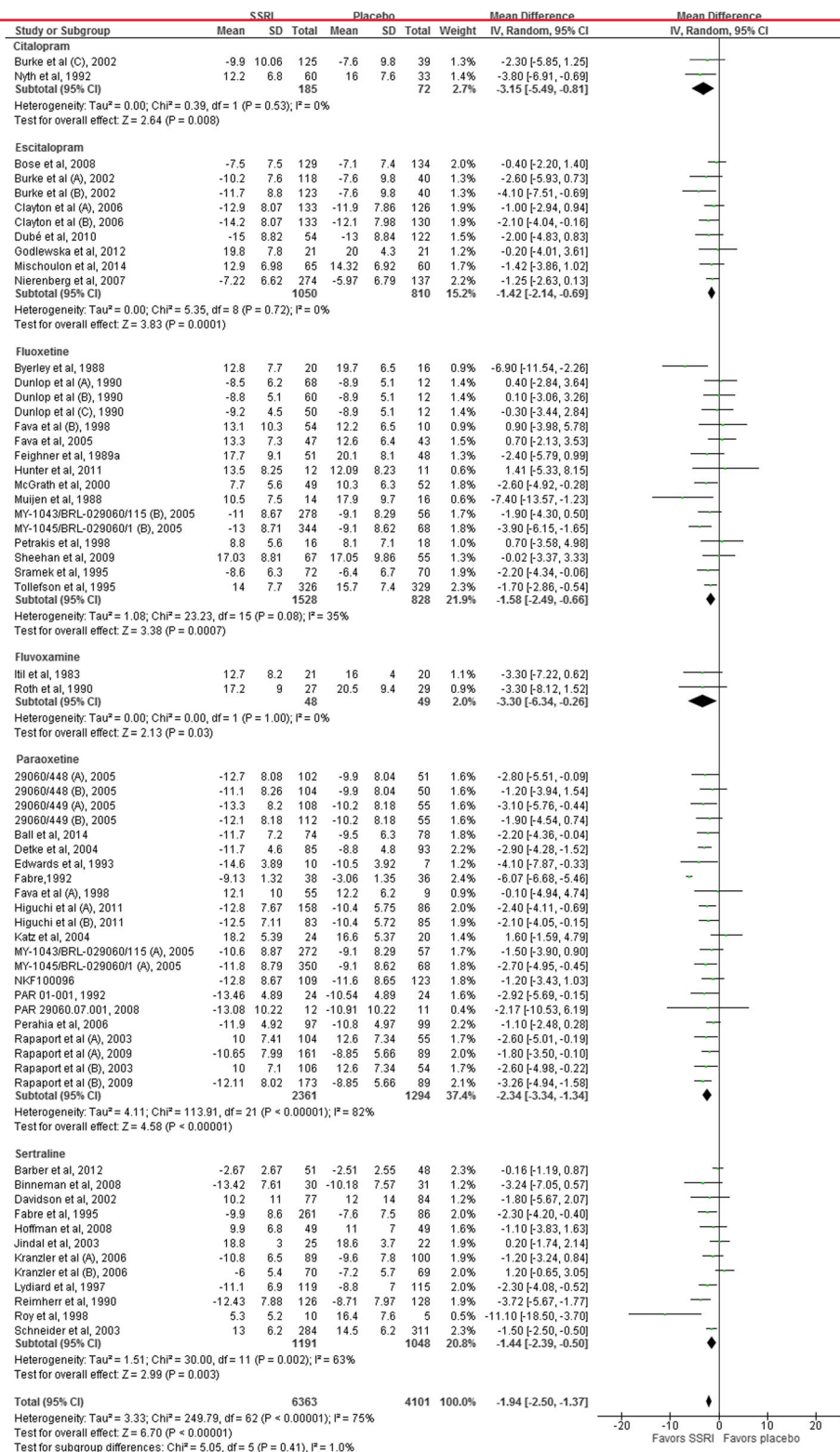
Jakobsen et al. BMC Psychiatry (2017) 17:58

Page 9 of 28

Study or Subgroup	SSRI Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Citalopram								
Burke et al (C), 2002	-9.9	10.06	125	-7.6	9.8	39	1.3%	-2.30 [-5.85, 1.25]
Nyth et al, 1992	12.2	6.8	60	16	7.6	33	1.4%	-3.80 [-6.91, -0.69]
Subtotal (95% CI)			185			72	2.7%	-3.15 [-5.49, -0.81]
Heterogeneity: Tau² = 0.00; Chi² = 0.39, df = 1 (P = 0.53); I² = 0%								
Test for overall effect: Z = 2.64 (P = 0.008)								
Escitalopram								
Bose et al, 2008	-7.5	7.5	129	-7.1	7.4	134	2.0%	-0.40 [-2.20, 1.40]
Burke et al (A), 2002	-10.2	7.6	118	-7.6	9.8	40	1.3%	-2.60 [-5.93, 0.73]
Burke et al (B), 2002	-11.7	8.8	123	-7.6	9.8	40	1.3%	-4.10 [-7.51, -0.69]
Clayton et al (A), 2006	-12.9	8.07	133	-11.9	7.86	126	1.9%	-1.00 [-2.94, 0.94]
Clayton et al (B), 2006	-14.2	8.07	133	-12.1	7.98	130	1.9%	-2.10 [-4.04, -0.16]
Dubé et al, 2010	-15	8.82	54	-13	8.84	122	1.5%	-2.00 [-4.83, 0.83]
Godlewska et al, 2012	19.8	7.8	21	20	4.3	21	1.2%	-0.20 [-4.01, 3.61]
Mischoulon et al, 2014	12.9	6.98	65	14.32	6.92	60	1.7%	-1.42 [-3.86, 1.02]
Nierenberg et al, 2007	-7.22	6.62	274	-5.97	6.79	137	2.2%	-1.25 [-2.63, 0.13]
Subtotal (95% CI)			1050			810	15.2%	-1.42 [-2.14, -0.69]
Heterogeneity: Tau² = 0.00; Chi² = 5.35, df = 8 (P = 0.72); I² = 0%								
Test for overall effect: Z = 3.83 (P = 0.0001)								
Fluoxetine								
Byerley et al, 1988	12.8	7.7	20	19.7	6.5	16	0.9%	-6.90 [-11.54, -2.26]
Dunlop et al (A), 1990	-8.5	6.2	68	-8.9	5.1	12	1.4%	0.40 [-2.84, 3.64]
Dunlop et al (B), 1990	-8.8	5.1	60	-8.9	5.1	12	1.4%	0.10 [-3.06, 3.26]
Dunlop et al (C), 1990	-9.2	4.5	50	-8.9	5.1	12	1.4%	-0.30 [-3.44, 2.84]
Fava et al (B), 1998	13.1	10.3	54	12.2	6.5	10	0.9%	0.90 [-3.98, 5.78]
Fava et al, 2005	13.3	7.3	47	12.6	6.4	43	1.5%	0.70 [-2.13, 3.53]
Feighner et al, 1989a	17.7	9.1	51	20.1	8.1	48	1.3%	-2.40 [-5.79, 0.99]
Hunter et al, 2011	13.5	8.25	12	12.09	8.23	11	0.6%	1.41 [-5.33, 8.15]
McGrath et al, 2000	7.7	5.6	49	10.3	6.3	52	1.8%	-2.60 [-4.92, -0.28]
Muijen et al, 1988	10.5	7.5	14	17.9	9.7	16	0.6%	-7.40 [-13.57, -1.23]
MY-1043/BRL-029060/115 (B), 2005	-11	8.67	278	-9.1	8.29	56	1.7%	-1.90 [-4.30, 0.50]
MY-1045/BRL-029060/1 (B), 2005	-13	8.71	344	-9.1	8.62	68	1.8%	-3.90 [-6.15, -1.65]
Petrakis et al, 1998	8.8	5.6	16	8.1	7.1	18	1.0%	0.70 [-3.58, 4.98]
Sheehan et al, 2009	17.03	8.81	67	17.05	9.86	55	1.3%	-0.02 [-3.37, 3.33]
Sramek et al, 1995	-8.6	6.3	72	-6.4	6.7	70	1.9%	-2.20 [-4.34, -0.06]
Tollefson et al, 1995	14	7.7	326	15.7	7.4	329	2.3%	-1.70 [-2.86, -0.54]
Subtotal (95% CI)			1528			828	21.9%	-1.58 [-2.49, -0.66]
Heterogeneity: Tau² = 1.08; Chi² = 23.23, df = 15 (P = 0.08); I² = 35%								
Test for overall effect: Z = 3.38 (P = 0.0007)								
Fluvoxamine								
Itil et al, 1983	12.7	8.2	21	16	4	20	1.1%	-3.30 [-7.22, 0.62]
Roth et al, 1990	17.2	9	27	20.5	9.4	29	0.9%	-3.30 [-8.12, 1.52]
Subtotal (95% CI)			48			49	2.0%	-3.30 [-6.34, -0.26]
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 1.00); I² = 0%								
Test for overall effect: Z = 2.13 (P = 0.03)								
Paraoxetine								
29060/448 (A), 2005	-12.7	8.08	102	-9.9	8.04	51	1.6%	-2.80 [-5.51, -0.09]
29060/448 (B), 2005	-11.1	8.26	104	-9.9	8.04	50	1.6%	-1.20 [-3.94, 1.54]
29060/449 (A), 2005	-13.3	8.2	108	-10.2	8.18	55	1.6%	-3.10 [-5.76, -0.44]
29060/449 (B), 2005	-12.1	8.18	112	-10.2	8.18	55	1.6%	-1.90 [-4.54, 0.74]
Ball et al, 2014	-11.7	7.2	74	-9.5	6.3	78	1.8%	-2.20 [-4.36, -0.04]
Detke et al, 2004	-11.7	4.6	85	-8.8	4.8	93	2.2%	-2.90 [-4.28, -1.52]
Edwards et al, 1993	-14.6	3.89	10	-10.5	3.92	7	1.2%	-4.10 [-7.87, -0.33]
Fabre, 1992	-9.13	1.32	38	-3.06	1.35	36	2.4%	-6.07 [-6.68, -5.46]
Fava et al (A), 1998	12.1	10	55	12.2	6.2	9	0.9%	-0.10 [-4.94, 4.74]
Higuchi et al (A), 2011	-12.8	7.67	158	-10.4	5.75	86	2.0%	-2.40 [-4.11, -0.69]
Higuchi et al (B), 2011	-12.5	7.11	83	-10.4	5.72	85	1.9%	-2.10 [-4.05, -0.15]
Katz et al, 2004	18.2	5.39	24	16.6	5.37	20	1.4%	1.60 [-1.59, 4.79]
MY-1043/BRL-029060/115 (A), 2005	-10.6	8.87	272	-9.1	8.29	57	1.7%	-1.50 [-3.90, 0.90]
MY-1045/029060/1 (A), 2005	-11.8	8.79	350	-9.1	8.62	68	1.8%	-2.70 [-4.95, -0.45]
NKF100096	-12.8	8.67	109	-11.6	8.65	123	1.8%	-1.20 [-3.43, 1.03]
PAR 01-001, 1992	-13.46	4.89	24	-10.54	4.89	24	1.6%	-2.92 [-5.69, -0.15]
PAR 29060.07.001, 2008	-13.08	10.22	12	-10.91	10.22	11	0.4%	-2.17 [-10.53, 6.19]
Perahia et al, 2006	-11.9	4.92	97	-10.8	4.97	99	2.2%	-1.10 [-2.48, 0.28]
Rapaport et al (A), 2003	10	7.41	104	12.6	7.34	55	1.7%	-2.60 [-5.01, -0.19]
Rapaport et al (A), 2009	-10.65	7.99	161	-8.85	5.66	89	2.0%	-1.80 [-3.50, -0.10]
Rapaport et al (B), 2003	10	7.1	106	12.6	7.34	54	1.7%	-2.60 [-4.98, -0.22]
Rapaport et al (B), 2009	-12.11	8.02	173	-8.85	5.66	89	2.1%	-3.26 [-4.94, -1.58]
Subtotal (95% CI)			2361			1294	37.4%	-2.34 [-3.34, -1.34]
Heterogeneity: Tau² = 4.11; Chi² = 113.91, df = 21 (P < 0.00001); I² = 82%								
Test for overall effect: Z = 4.58 (P < 0.00001)								
Sertraline								
Barber et al, 2012	-2.67	2.67	51	-2.51	2.55	48	2.3%	-0.16 [-1.19, 0.87]
Binneman et al, 2008	-13.42	7.61	30	-10.18	7.57	31	1.2%	-3.24 [-7.05, 0.57]
Davidson et al, 2002	10.2	11	77	12	14	84	1.2%	-1.80 [-5.67, 2.07]
Fabre et al, 1995	-9.9	8.6	261	-7.6	7.5	86	2.0%	-2.30 [-4.20, -0.40]
Hoffman et al, 2008	9.9	6.8	49	11	7	49	1.6%	-1.10 [-3.83, 1.63]
Jindal et al, 2003	18.8	3	25	18.6	3.7	22	1.9%	0.20 [-1.74, 2.14]
Kranzler et al (A), 2006	-10.8	6.5	89	-9.6	7.8	100	1.9%	-1.20 [-3.24, 0.84]
Kranzler et al (B), 2006	-6	5.4	70	-7.2	5.7	69	2.0%	1.20 [-0.65, 3.05]
Lydiard et al, 1997	-11.1	6.9	119	-8.8	7	115	2.0%	-2.30 [-4.08, -0.52]
Reimherr et al, 1990	-12.43	7.88	126	-8.71	7.97	128	1.9%	-3.72 [-5.67, -1.77]
Roy et al, 1998	5.3	5.2	10	16.4	7.6	5	0.5%	-11.10 [-18.50, -3.70]
Schneider et al, 2003	13	6.2	284	14.5	6.2	311	2.3%	-1.50 [-2.50, -0.50]
Subtotal (95% CI)			1191			1048	20.8%	-1.44 [-2.39, -0.50]
Heterogeneity: Tau² = 1.51; Chi² = 30.00, df = 11 (P = 0.002); I² = 63%								
Test for overall effect: Z = 2.99 (P = 0.003)								
Total (95% CI)			6363			4101	100.0%	-1.94 [-2.50, -1.37]
Heterogeneity: Tau² = 3.33; Chi² = 249.79, df = 62 (P < 0.00001); I² = 75%								
Test for overall effect: Z = 6.70 (P < 0.00001)								
Test for subgroup differences: Chi² = 5.05, df = 5 (P = 0.41), I² = 1.0%								

Favors SSRI Favors placebo

Fig. 4 Subgroup analysis comparing trials assessing the effects of different selective serotonin reuptake inhibitors

Jakobsen et al. BMC Psychiatry (2017) 17:58
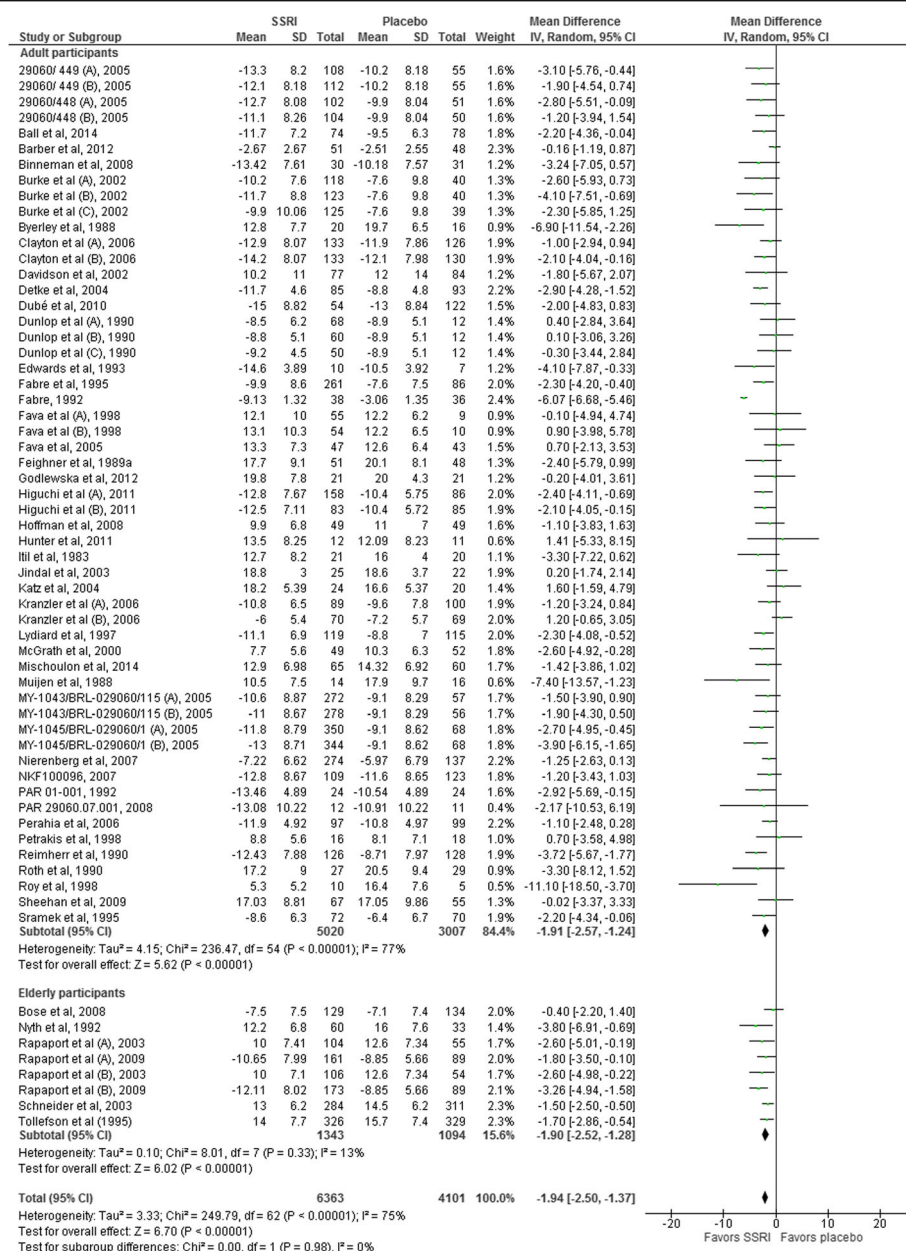
Page 10 of 28

Study or Subgroup	SSRI Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Adult participants								
29060/ 449 (A), 2005	-13.3	8.2	108	-10.2	8.18	55	1.6%	-3.10 [-5.76, -0.44]
29060/ 449 (B), 2005	-12.1	8.18	112	-10.2	8.18	55	1.6%	-1.90 [-4.54, 0.74]
29060/448 (A), 2005	-12.7	8.08	102	-9.9	8.04	51	1.6%	-2.80 [-5.51, -0.09]
29060/448 (B), 2005	-11.1	8.26	104	-9.9	8.04	50	1.6%	-1.20 [-3.94, 1.54]
Ball et al, 2014	-11.7	7.2	74	-9.5	6.3	78	1.8%	-2.20 [-4.36, -0.04]
Barber et al, 2012	-2.67	2.67	51	-2.51	2.55	48	2.3%	-0.16 [-1.19, 0.87]
Binneman et al, 2008	-13.42	7.61	30	-10.18	7.57	31	1.2%	-3.24 [-7.05, 0.57]
Burke et al (A), 2002	-10.2	7.6	118	-7.6	9.8	40	1.3%	-2.60 [-5.93, 0.73]
Burke et al (B), 2002	-11.7	8.8	123	-7.6	9.8	40	1.3%	-4.10 [-7.51, -0.69]
Burke et al (C), 2002	-9.9	10.06	125	-7.6	9.8	39	1.3%	-2.30 [-5.85, 1.25]
Byerley et al, 1988	12.8	7.7	20	19.7	6.5	16	0.9%	-6.90 [-11.54, -2.26]
Clayton et al (A), 2006	-12.9	8.07	133	-11.9	7.86	126	1.9%	-1.00 [-2.94, 0.94]
Clayton et al (B), 2006	-14.2	8.07	133	-12.1	7.98	130	1.9%	-2.10 [-4.04, -0.16]
Davidson et al, 2002	10.2	11	77	12	14	84	1.2%	-1.80 [-5.67, 2.07]
Detke et al, 2004	-11.7	4.6	85	-8.8	4.8	93	2.2%	-2.90 [-4.28, -1.52]
Dubé et al, 2010	-15	8.82	54	-13	8.84	122	1.5%	-2.00 [-4.83, 0.83]
Dunlop et al (A), 1990	-8.5	6.2	68	-8.9	5.1	12	1.4%	0.40 [-2.84, 3.64]
Dunlop et al (B), 1990	-8.8	5.1	60	-8.9	5.1	12	1.4%	0.10 [-3.06, 3.26]
Dunlop et al (C), 1990	-9.2	4.5	50	-8.9	5.1	12	1.4%	-0.30 [-3.44, 2.84]
Edwards et al, 1993	-14.6	3.89	10	-10.5	3.92	7	1.2%	-4.10 [-7.87, -0.33]
Fabre et al, 1995	-9.9	8.6	261	-7.6	7.5	86	2.0%	-2.30 [-4.20, -0.40]
Fabre, 1992	-9.13	1.32	38	-3.06	1.35	36	2.4%	-6.07 [-6.68, -5.46]
Fava et al (A), 1998	12.1	10	55	12.2	6.2	9	0.9%	-0.10 [-4.94, 4.74]
Fava et al (B), 1998	13.1	10.3	54	12.2	6.5	10	0.9%	0.90 [-3.98, 5.78]
Fava et al, 2005	13.3	7.3	47	12.6	6.4	43	1.5%	0.70 [-2.13, 3.53]
Feighner et al, 1989a	17.7	9.1	51	20.1	8.1	48	1.3%	-2.40 [-5.79, 0.99]
Godlewska et al, 2012	19.8	7.8	21	20	4.3	21	1.2%	-0.20 [-4.01, 3.61]
Higuchi et al (A), 2011	-12.8	7.67	158	-10.4	5.75	86	2.0%	-2.40 [-4.11, -0.69]
Higuchi et al (B), 2011	-12.5	7.11	83	-10.4	5.72	85	1.9%	-2.10 [-4.05, -0.15]
Hoffman et al, 2008	9.9	6.8	49	11	7	49	1.6%	-1.10 [-3.83, 1.63]
Hunter et al, 2011	13.5	8.25	12	12.09	8.23	11	0.6%	1.41 [-5.33, 8.15]
Itil et al, 1983	12.7	8.2	21	16	4	20	1.1%	-3.30 [-7.22, 0.62]
Jindal et al, 2003	18.8	3	25	18.6	3.7	22	1.9%	0.20 [-1.74, 2.14]
Katz et al, 2004	18.2	5.39	24	16.6	5.37	20	1.4%	1.60 [-1.59, 4.79]
Kranzler et al (A), 2006	-10.8	6.5	89	-9.6	7.8	100	1.9%	-1.20 [-3.24, 0.84]
Kranzler et al (B), 2006	-6	5.4	70	-7.2	5.7	69	2.0%	1.20 [-0.65, 3.05]
Lydiard et al, 1997	-11.1	6.9	119	-8.8	7	115	2.0%	-2.30 [-4.08, -0.52]
McGrath et al, 2000	7.7	5.6	49	10.3	6.3	52	1.8%	-2.60 [-4.92, -0.28]
Mischoulon et al, 2014	12.9	6.98	65	14.32	6.92	60	1.7%	-1.42 [-3.86, 1.02]
Muijen et al, 1988	10.5	7.5	14	17.9	9.7	16	0.6%	-7.40 [-13.57, -1.23]
MY-1043/BRL-029060/115 (A), 2005	-10.6	8.87	272	-9.1	8.29	57	1.7%	-1.50 [-3.90, 0.90]
MY-1043/BRL-029060/115 (B), 2005	-11	8.67	278	-9.1	8.29	56	1.7%	-1.90 [-4.30, 0.50]
MY-1045/BRL-029060/1 (A), 2005	-11.8	8.79	350	-9.1	8.62	68	1.8%	-2.70 [-4.95, -0.45]
MY-1045/BRL-029060/1 (B), 2005	-13	8.71	344	-9.1	8.62	68	1.8%	-3.90 [-6.15, -1.65]
Nierenberg et al, 2007	-7.22	6.62	274	-5.97	6.79	137	2.2%	-1.25 [-2.63, 0.13]
NKF100096, 2007	-12.8	8.67	109	-11.6	8.65	123	1.8%	-1.20 [-3.43, 1.03]
PAR 01-001, 1992	-13.46	4.89	24	-10.54	4.89	24	1.6%	-2.92 [-5.69, -0.15]
PAR 29060.07.001, 2008	-13.08	10.22	12	-10.91	10.22	11	0.4%	-2.17 [-10.53, 6.19]
Perahia et al, 2006	-11.9	4.92	97	-10.8	4.97	99	2.2%	-1.10 [-2.48, 0.28]
Petrakis et al, 1998	8.8	5.6	16	8.1	7.1	18	1.0%	0.70 [-3.58, 4.98]
Reimherr et al, 1990	-12.43	7.88	126	-8.71	7.97	128	1.9%	-3.72 [-5.67, -1.77]
Roth et al, 1990	17.2	9	27	20.5	9.4	29	0.9%	-3.30 [-8.12, 1.52]
Roy et al, 1998	5.3	5.2	10	16.4	7.6	5	0.5%	-11.10 [-18.50, -3.70]
Sheehan et al, 2009	17.03	8.81	67	17.05	9.86	55	1.3%	-0.02 [-3.37, 3.33]
Sramek et al, 1995	-8.6	6.3	72	-6.4	6.7	70	1.9%	-2.20 [-4.34, -0.06]
Subtotal (95% CI)			5020			3007	84.4%	-1.91 [-2.57, -1.24]

Heterogeneity: Tau² = 4.15; Chi² = 236.47, df = 54 (P < 0.00001); I² = 77%
Test for overall effect: Z = 5.62 (P < 0.00001)

Study or Subgroup	SSRI Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Elderly participants								
Bose et al, 2008	-7.5	7.5	129	-7.1	7.4	134	2.0%	-0.40 [-2.20, 1.40]
Nyth et al, 1992	12.2	6.8	60	16	7.6	33	1.4%	-3.80 [-6.91, -0.69]
Rapaport et al (A), 2003	10	7.41	104	12.6	7.34	55	1.7%	-2.60 [-5.01, -0.19]
Rapaport et al (A), 2009	-10.65	7.99	161	-8.85	5.66	89	2.0%	-1.80 [-3.50, -0.10]
Rapaport et al (B), 2003	10	7.1	106	12.6	7.34	54	1.7%	-2.60 [-4.98, -0.22]
Rapaport et al (B), 2009	-12.11	8.02	173	-8.85	5.66	89	2.1%	-3.26 [-4.94, -1.58]
Schneider et al, 2003	13	6.2	284	14.5	6.2	311	2.3%	-1.50 [-2.50, -0.50]
Tollefson et al (1995)	14	7.7	326	15.7	7.4	329	2.3%	-1.70 [-2.86, -0.54]
Subtotal (95% CI)			1343			1094	15.6%	-1.90 [-2.52, -1.28]

Heterogeneity: Tau² = 0.10; Chi² = 8.01, df = 7 (P = 0.33); I² = 13%
Test for overall effect: Z = 6.02 (P < 0.00001)

Total (95% CI)			6363			4101	100.0%	-1.94 [-2.50, -1.37]

Heterogeneity: Tau² = 3.33; Chi² = 249.79, df = 62 (P < 0.00001); I² = 75%
Test for overall effect: Z = 6.70 (P < 0.00001)
Test for subgroup differences: Chi² = 0.00, df = 1 (P = 0.98), I² = 0%

−20 −10 0 10 20
Favors SSRI Favors placebo

Fig. 5 Subgroup analysis comparing trials randomising elderly participants to trials randomising non-elderly participants

SSRI group and the placebo group (−0.92 points; 95% CI −2.42 to 0.58; I² 26%) (Additional file 4: Figure S1).

Trial Sequential Analysis
The required information size was calculated based on an anticipated intervention effect of 3 HDRS points, the empirical variance, a risk of type I error of 0.025, and a power of 80% [12, 13, 17]. The Trial Sequential Analysis showed that the trial monitoring boundary for benefit was crossed after the 9th trial indicating a statistically significant result (Trial Sequential Analysis-adjusted CI −2.62 to −1.26) (Fig. 7).

Bayes factor
Bayes factor was calculated based on our anticipated intervention effect of 3 HDRS points and the primary meta-analysis result (mean difference −1.94 points; 95% CI −2.50 to −1.37) [12, 13, 17]. Bayes factor ($2.01*10^{-23}$) was below the threshold for significance of 0.1, supporting the statistical significant result.

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 11 of 28
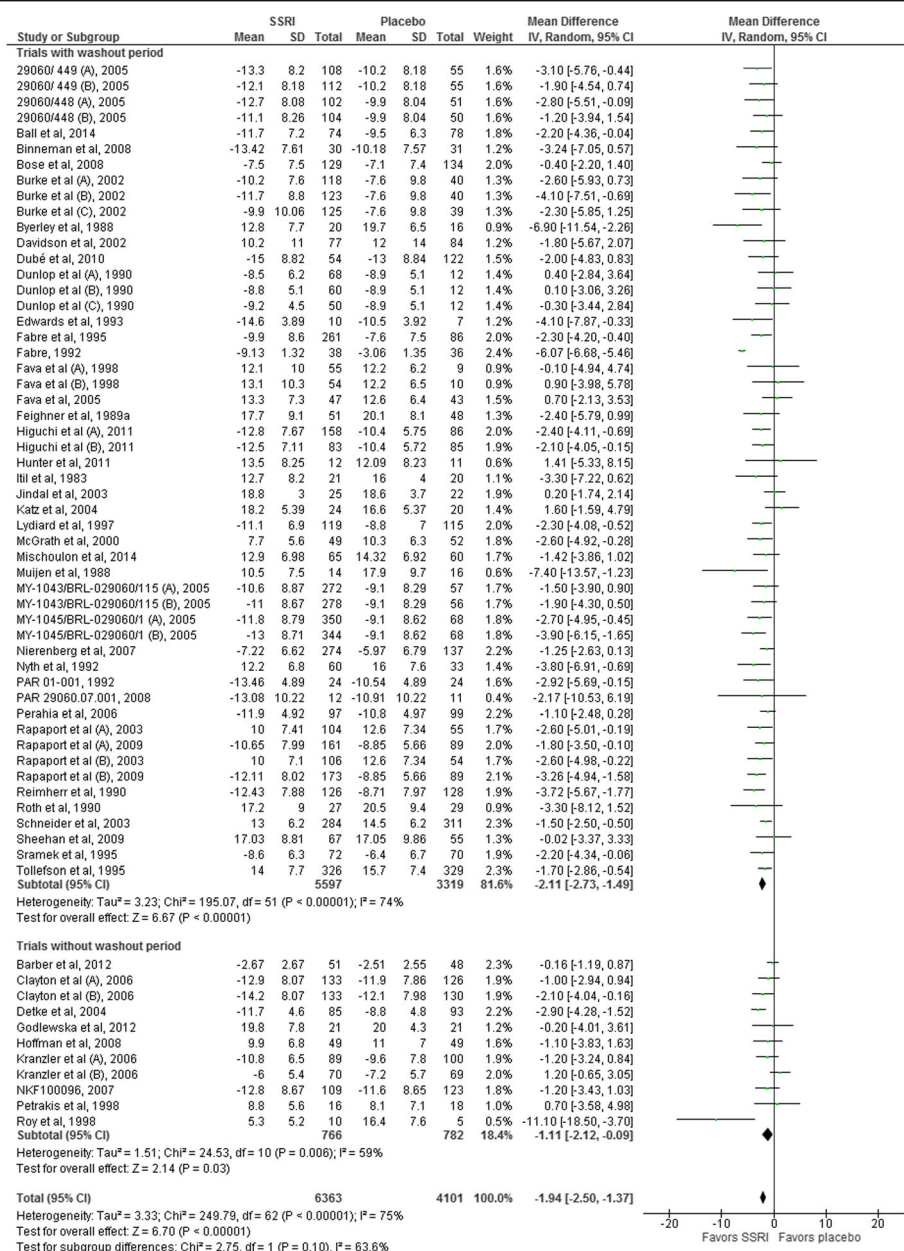
Fig. 6 Subgroup analysis comparing trials using a placebo-washout period to trials not using a placebo-washout period

Risk of publication bias

Visual inspection of the funnel plot did not show clear signs of asymmetry (Fig. 8) [13, 17].

Hamilton Depression Rating Scale, Montgomery-Asberg Depression Rating Scale, and Beck's Depression Inventory

Sixty-one trials [32–52, 54–66, 103–106, 120, 121, 123, 137–142, 144–146, 149–151, 154, 155, 168, 172, 173] reported means and SD on the HDRS [18], MADRS [19], or BDI [174]. Using standardised mean difference, random-effects meta-analysis showed that SSRI versus placebo significantly decreased the standardised mean

difference score (trials reporting mean scores: −0.23; 95% CI −0.31 to −0.14; *P* < 0.00001; trials reporting mean change scores: −0.26; 95% CI −0.35 to −0.17; *P* < 0.00001). The standardised mean difference was below our predefined threshold for clinical significance.

Long-term follow-up

One trial [33] reported mean HDRS scores and SD and one trial [146] reported mean HDRS change scores and SD at end of long-term follow-up. Random-effects meta-analysis of these trials showed a mean difference −0.18 points (95% CI −2.78 to 2.43; *P* = 0.89). Four trials

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 12 of 28
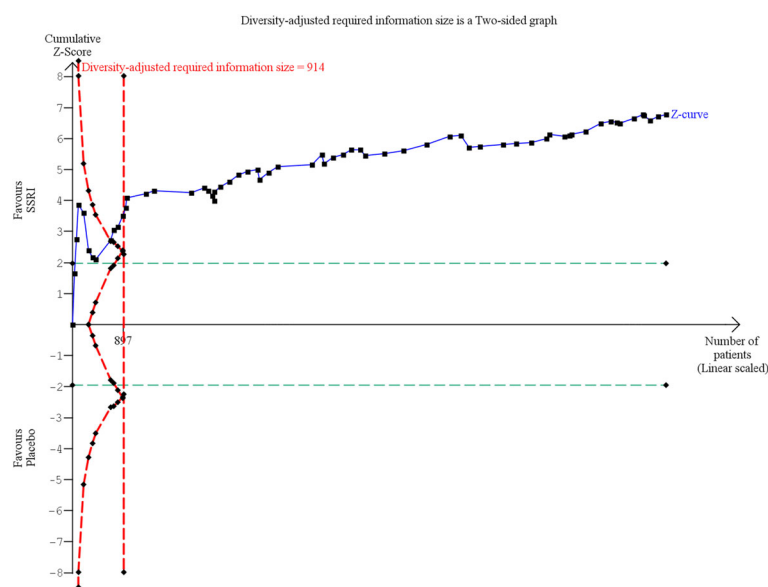
Fig. 7 Trial Sequential Analysis of the results of selective serotonin reuptake inhibitors on Hamilton depression rating scale

reported mean HDRS scores or change scores at end of long-term follow-up but without reporting SDs [70, 97, 107, 143]. SDs were imputed. Random-effects meta-analysis of all the six trials showed a mean difference of −1.30 points (95% CI −2.72 to 0.13; $P = 0.07$).

No remission

Thirty-four trials [33, 34, 38, 45, 49–53, 55, 56, 60, 78, 81, 101, 104, 107–112, 120, 128, 140, 141, 146, 153, 157, 158, 162, 171, 175, 176] reported the proportion of participants with no remission at end of treatment. A total of 1430/2211 (64.7%) SSRI participants experienced no remission compared with 1493/2003 (74.5%) control participants. Random-effects meta-analysis showed that SSRIs versus placebo significantly decreased the risk of no remission (RR 0.88; 95% CI 0.84 to 0.91; $P < 0.00001$) (Fig. 9). This corresponds to 657 (95% CI 642 to 679) SSRI participants out of 1000 will experience no remission compared with 746 control participants out of 1000 (see Additional file 5: Summary of findings table). Visual inspection of the forest plots showed no clear signs of heterogeneity [13, 16].

The required information size was calculated based on the observed proportion of control participants without remission, a relative risk reduction of 30%, a risk of type I error of 0.025, and a power of 80% [13, 17]. The Trial Sequential Analysis showed that the trial monitoring boundary for benefit was crossed (Fig. 10) and the Trial Sequential Analysis-adjusted CI was 0.83 to 0.92 [13, 17].

Bayes factor was 1426.8 based on the random-effects meta-analysis result and above the threshold for significance of 0.1 [13, 17]. This relatively high Bayes factor

indicates that it is more likely that null effect (null hypothesis) compared to the anticipated intervention effect has produced this meta-analysis results [12, 13].

Best-worst case scenario showed a highly significant meta-analysis result (RR 0.78; 95% CI 0.73 to 0.83; $P < 0.00001$) [13, 17]. Worst-best case scenario showed no significant difference on risk of no remission (RR 0.95; 95% CI 0.89 to 1.02; $P = 0.14$) [13, 17].

Visual inspection of the funnel plot showed no clear signs of publication bias [13, 16, 17].

Serious adverse events

Because of the low proportion in the control group (around 2%) we used the Mantel–Haenszel (MH) odds
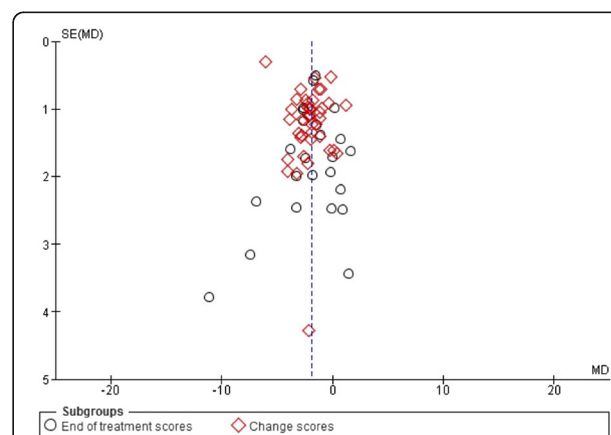
Fig. 8 Funnel plot of the random-effects meta-analysis of the effect of selective serotonin reuptake inhibitors on Hamilton depression rating scale

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 13 of 28
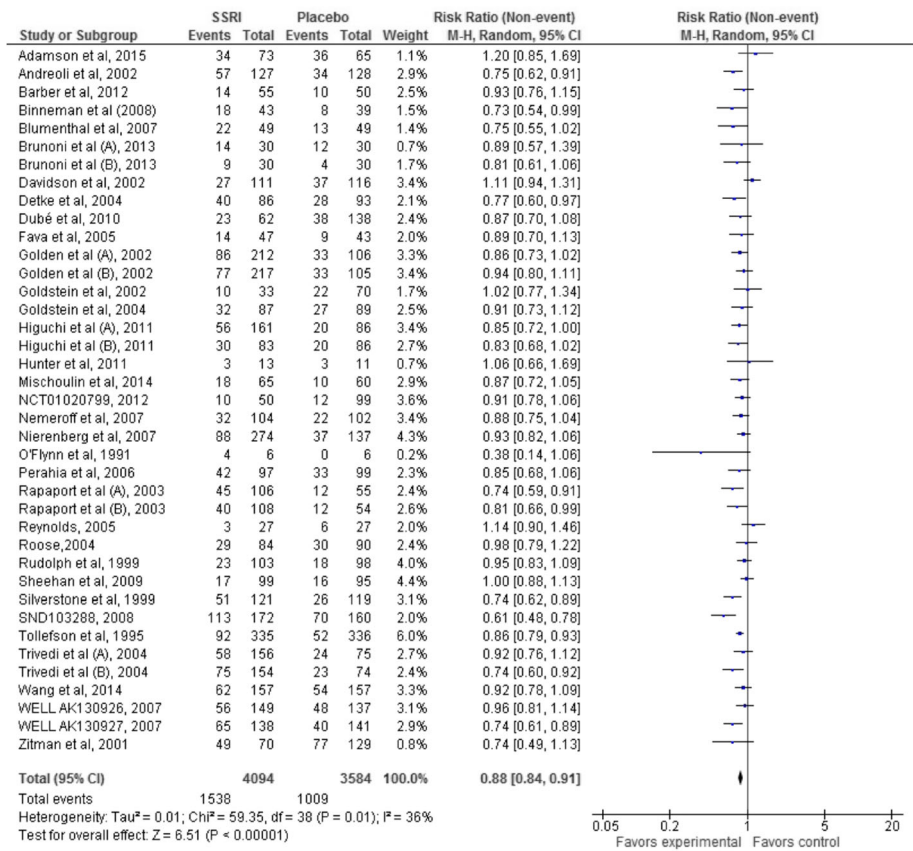
Fig. 9 Random-effects meta-analysis of the results of selective serotonin reuptake inhibitors on remission of depression

Fig. 10 Trial Sequential Analysis of the results of selective serotonin reuptake inhibitors on remission of depression

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 14 of 28
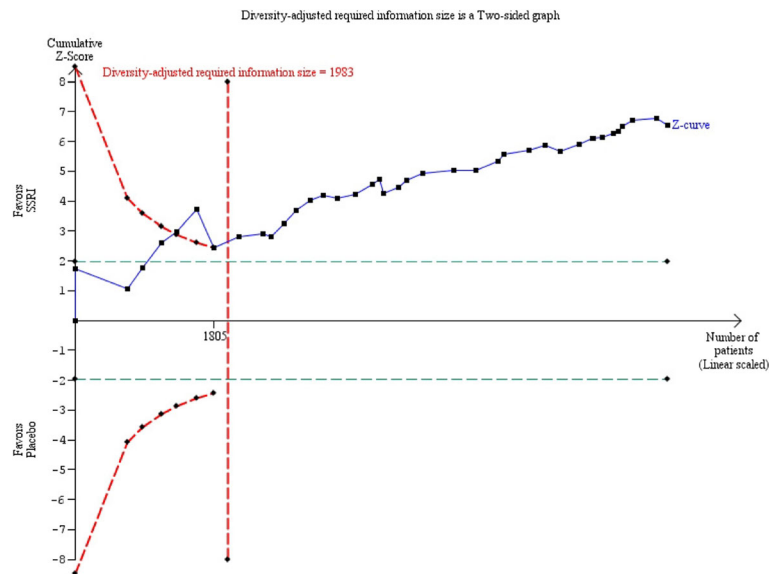
ratio method with reciprocal zero cell correction (zero is replaced by the reciprocal of the size of the opposite treatment arm) [177]. Forty-four trials reported the proportion of participants with serious adverse events [48, 49, 54–56, 60, 63, 75, 78, 93, 94, 102, 105, 108, 112–118, 120, 121, 137–139, 141, 144–146, 149–151, 155, 167, 171, 173, 176, 178–180]. A total of 239/8242 (2.7%) SSRI participants experienced a serious adverse event compared with 106/4956 (2.1%) control participants. Random-effects meta-analysis showed that SSRIs versus placebo significantly increased the risk of a serious adverse event (OR 1.37; 95% CI 1.08 to 1.75; P = 0.009) (Fig. 11). This corresponds to 31 (95% CI 25 to 40)/1000 SSRI participants will experience a serious adverse event compared with 22/1000 control participants (see Additional file 5: Summary of findings table). Visual inspection of the forest plot did not indicate significant heterogeneity [13, 17]. Even when using the multiplicity adjusted risk of type I error (0.05/2 =

0.025), the trial sequential boundary for harm was crossed (Fig. 12) and the Trial Sequential Analysis-adjusted CI was 1.03 to 1.89 [13, 17]. Table 2 summarises the types of adverse events.

Bayes factor was $4.8*10^5$ above the threshold for significance of 0.1 [13, 17]. This clearly shows that a beneficial effect of SSRIs on serious adverse events is very unlikely [13, 17]. Visual inspection of the funnel plot showed no signs of publication bias [13, 16, 17]. Based on the random-effects meta-analysis result, we calculated the number-needed-to-seriously harm one patient to be 138 patients.

Adverse events

Meta-analyses showed that the participants randomised to SSRIs versus placebo had a significantly increased risk of several adverse events. We have summarised the risks of the adverse events which were most reported in
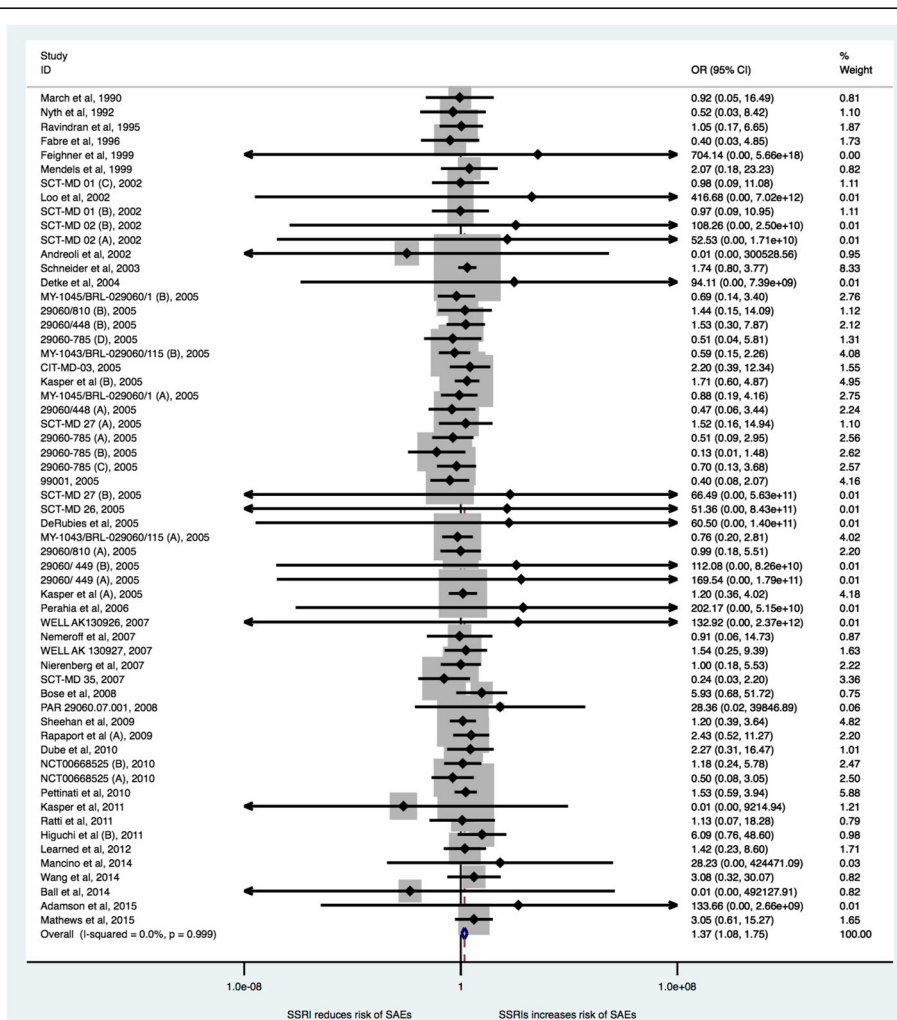
Fig. 11 Random-effects meta-analysis of the results of selective serotonin reuptake inhibitors on serious adverse events
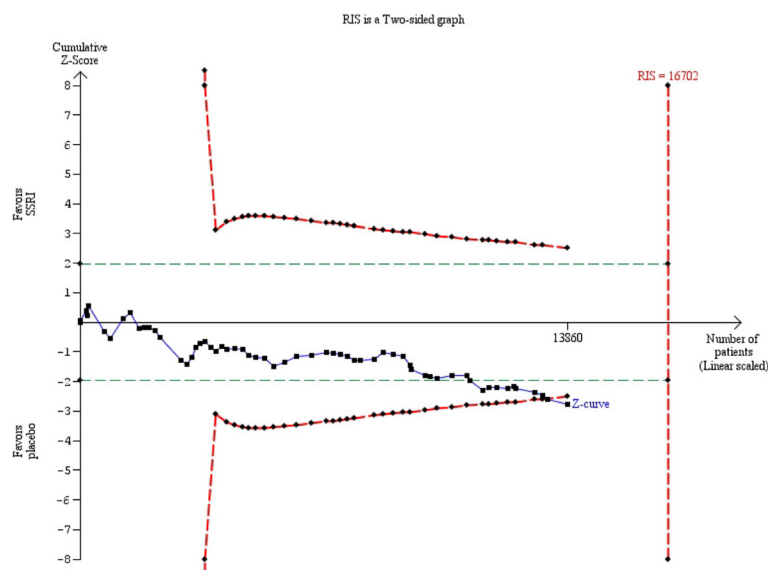
Fig. 12 Trial Sequential Analysis of the results of selective serotonin reuptake inhibitors on serious adverse events

Additional file 6, including numbers-needed-to-harm. We have also included forest plots for the 25 most statistically significant adverse event results in the Additional files (see Additional files 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 and 31: Figure S3-S27). The full list of the 84 reported adverse events are summarised in Table 3.

Clinical significance

All primary HDRS meta-analyses showed intervention effect estimates below our predefined threshold for clinical significance (a mean difference of 3 HDRS points or 0.5 standardised mean difference) [13, 17]. Our results show statistically significant effects, but the possible effects all seem to have questionable clinical significance [13].

Secondary outcomes
Suicides, suicide attempts, and suicide ideation

There were no significant differences between participants randomised to SSRIs versus placebo on number of suicides (RR 0.68; 95% CI 0.16 to 2.81; $P = 0.59$; Trial Sequential Analysis-adjusted CI 0.01 to 226.85; 6 trials [60, 71, 108, 113, 151, 155]); suicide attempts (RR 1.76; 95% CI 0.59 to 5.22; $P = 0.31$; Trial Sequential Analysis-adjusted CI 0.02 to 149.95; 8 trials [49, 56, 75, 94, 102, 139, 167, 181]); or suicide ideation (RR 0.80; 95% CI 0.36 to 1.77; $P = 0.58$; Trial Sequential Analysis-adjusted CI 0.03 to 23.20; 11 trials [49, 51, 120, 138, 139, 145, 151, 162, 167, 171, 180]). The required information size was not reached in any of the three Trial Sequential Analyses.

Quality of life

Only six trials assessed quality of life [48, 51, 63, 100, 101, 112] out of which four trials reported results on the quality of life enjoyment and satisfaction questionnaire (Q-LES-Q) [48, 51, 100, 101]. Two trials [48, 101] reported mean scores and SDs. Random-effects meta-analysis showed significant effect of SSRI on Q-LES-Q scores (RR 2.98; 95% CI 1.34 to 4.61; $P = 0.0004$).

Two trials reported results on the short form of the quality of life enjoyment and satisfaction questionnaire [63, 112], but only one trial reported mean scores and SDs [63]. The results from this trial showed that SSRIs (paroxetine) versus placebo significantly increased the mean score of the questionnaire (paroxetine 12.5 mg group mean 11.4, SD 16.7; paroxetine 25 mg group mean 11.5, SD 17.2; placebo group mean 5.3, SD 17.1) [63].

Post hoc analysis of no response

We identified 70 trials assessing the effects of SSRIs on no response defined as less than 50% reduction (from baseline) on either HDRS or MADRS. The meta-analysis showed that SSRIs seem to significantly decrease the risk of no response compared with placebo (RR 0.83; 95% CI 0.80 to 0.87; $P = 0.00001$) (Additional file 32: Figure S2).

GRADE assessments

GRADE assessments show that due to the high risks of bias the quality of the evidence must be regarded as very low (Additional file 5: Summary of Findings Table) [13].

Discussion

SSRIs may affect the concentration of essential neurotransmitter substances in the brain and are therefore

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 16 of 28

Table 2 Summary of serious adverse events in the included trials

Trial	Experimental intervention	SSRI participants assessed for serious adverse events		Placebo participants assessed for serious adverse events	
		Numbers and types of serious adverse events	Proportion of participants with a serious adverse event	Numbers and types of serious adverse events	Proportion of participants with a serious adverse event
Bose et al., 2008	Escitalopram	1 bowel obstruction, 1 nausea, 1 arrythmia, 1 respiratory arrest, 1 retinal detachment, 1 chest pain	5 out of 96	1 syncope	1 out of 109
Ball et al., 2014	Paroxetine	No serious adverse event	0 out of 74	1 unspecified serious adverse event	1 out of 78
Andreoli et al., 2002	Fluoxetine	No serious adverse event	0 out of 97	1 suicide	1 out of 76
Kasper et al., 2011	Escitalopram	No serious adverse event	0 out of 128	1 hospitalisation due to appendicitis	1 out of 50
Kasper et al. (A), 2005	Escitalopram	1 death (suicide)	1 out of 173	1 death (probably drowned)	1 out of 160
Ravindran et al., 1995	Sertraline	9 unspecified serious adverse events	4 out of 25	3 unspecified serious adverse events	2 out of 13
March et al., 1990	Fluvoxamine	1 hospitalisation due to worsening of depression	1 out of 13	1 suicide attempt	1 out of 12
Rapaport et al., 2009	Paroxetine	2 chestpain, 1 osteoarthritis, 1 ankle fracture, 1 atrial fibrillation, 1 femur fracture,1 coronary artery occlusion, 1 pneumonia, 1 confusional state, 1 depression	10 out of 267	1 neprholithiasis, 1 aortic aneurism	2 out of 127
Higuchi et al., 2011	Paroxetine	1 suicide and 8 unspecified serious adverse events	9 out of 213	1 unspecified serious adverse event	1 out of 139
Schneider et al., 2003	Sertraline	17 unspecified serious adverse events	17 out of 284	11 unspecified serious adverse events	11 out of 311
Sheehan et al., 2009	Fluoxetine	1 suicidal ideation/suicidality, 1 worsening of depression, 2 suicide attempts, 1 anxiety/agitation/racing thoughts, 1 syncope, 1 ankle fracture, 1 viral gastro enteritis	8 out of 76	2 suicidal ideation/suicidality, 2 worsening of depression, 1 nose bleed, 1 allergic reaction	6 out of 67
Nemroff et al., 2005	Fluxetine	1 unspecified serious adverese event	1 out of 86	1 unspecified serious adverse event	1 out of 78
Fabre et al., 1996	Fluvoxamine	1 hospitalisation (non-cardiac chest pain)	1 out of 22	1 hospitalisation, 1 ruptured ectopic pregnancy, 1 hernia repair	2 out of 19
Feighner et al., 1999	Citalopram	3 suicide attempts, 1 miscarriage, 1 intestinal flu symptoms, 1 chest pain, 1 severe thinking abnormality, 1 allergic reaction	8 out of 349	No serious adverse event	0 out of 86
SCT-MD 01 (B), 2002	Escitalopram	1 anaphylaxis, 1 suicide attempt	2 out of 94	1 gallbladder stones	1 out of 46
SCT-MD 01 (C), 2002	Citalopram	1 coma, 1 intestinal fistula	2 out of 93	1 non-accidental overdose	1 out of 45
SCT-MD 02 (A), 2002	Escitalopram	1 suicidal tendency, suicide attempt; 1 non-accidental overdose, suidal attempt, tachycardia	2 out of 96	No serious adverse event	0 out of 53
SCT-MD 02 (B), 2002	Citalopram	1 cholestasis intrahepatic, dehydration	1 out of 99	No serious adverse event	0 out of 52
Pettinati et al., 2010	Sertraline	15 unspecified serious adverse events	15 out of 40	11 unspecified serious adverse events	11 out of 39
Dube et al., 2010	Escitalopram	1 suicide attempt, 1 gastro enteritis/ malaria	2 out of 54	1 near drowning, 1 gastro enteritis	2 out of 122
Learned et al., 2012	Paroxetine	1 intentional over dose, 1 depression, 1 unspecified event	3 out of 166	2 unspecified serious adverese events	2 out of 156

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 17 of 28

Table 2 Summary of serious adverse events in the included trials *(Continued)*

Ratti et al., 2011	Paroxetine	1 hemorrhoidal hemorrhage	1 out of 109	1 rash	1 out of 123
Wang et al., 2014	Escitalopram	3 unspecified serious adverese event	3 out of 114	1 unspecified serious adverese event	1 out of 115
Detke et al., 2004	Paroxetine	1 unspecified serious adverse event	1 out of 85	No serious adverse event	0 out of 93
Mancino et al., 2014	Sertraline	1 hospitalization	1 out of 23	No serious adverse event	0 out of 27
DeRubeis et al., 2005	Paroxetine	1 suicide	1 out of 120	No serious adverse event	0 out of 60
29060-785 (A), 2005	Paroxetine CR 25 mg	3 abnormal laboratory vaue, 1 emotional lability	4 out of 98	1 abnormal laboratory value, 1 gastrointestinal disorder	2 out of 26
29060-785 (B), 2005	Paroxetine CR 12.5 mg	1 abnormal laboratory value	1 out of 94	1 abnormal laboratory value, 1 myocardial infarction	2 out of 26
29060-785 (C), 2005	Citalopram 20 mg	5 abnormal laboratory value, 1 syncope	6 out of 105	1 abnormal laboratory value, 1 suicide	2 out of 25
29060-785 (D), 2005	Citalopram 40 mg	1 abnormal laboratory vaue, 1 emotional lability	2 out of 97	1 abnormal laboratory value	1 out of 25
SCT-MD 27 (A), 2005	Escitalopram	1 depression, 1 abnormal mental status, 1 malignant neoplasm	3 out of 131	1 labyrinthitis	1 out of 66
SCT-MD 27 (B), 2005	Sertraline	1 appendicitis	1 out of 135	No serious adverse event	0 out of 66
SCT-MD 35, 2007	Escitalopram	1 abnormal hepatic function	1 out of 131	1 breast cancer, 1 depression, 1 suicidal ideation, 1 suicide	4 out of 130
SCT-MD 26, 2005	Escitalopram	1 inflicted injury	1 out of 143	No serious adverse event	0 out of 151
MY-1043/BRL-029060/115 (A), 2005	Paroxetine	1 hypertension, 1 diabetes and hypothyroidism, 1 Fibrocystic disease, 1 Ovarian cysts, 1 peptic ulcer hemorrhage, 1 spinal surgery, 1 hypomanic episode with suicidal tendency, 2 Suicidal ideation, 1 alcoholism, 1 neoplasm	11 out of 272	1 suicidal ideation, 1 back pain, 1 trauma	3 out of 57
MY-1043/BRL-029060/115 (B)	Fluoxetine	1 suicidal ideation, 1 neoplasm, 2 acute pyelonephritis, 1 thrombophlebitis, 1 ectopic pregnancy, 1 polycystic granuloma, 2 basal cell carcinomas, 1 myxoid mitral valve	9 out of 278 (10 SAE in 9 participants)	1 viral meningitis, 1 infection, 1 myocardial infarction, 1 mole removal	3 out of 56 (4 SAE in 3 participants)
MY-1045/BRL-029060/1, (A), 2005	Paroxetine	2 depression (worsening), 2 emotional lability, 1 neoplasm, 1 insomnia, 1 nervousness, 1 carcinoma, 1 epistaxis, 1 gastro intestinal disorder, 1 prostate disorder	9 out of 357 (11 SAE in 9 participants)	1 depression (worsening), 1 rectal disorder	2 out of 70
MY-1045/BRL-029060/1, 2005	Fluoxetine	1 depression (worsening), 2 emotional lability, 1 neo plasm, 1 coronary artery disease, 1 thrombo phlebitis, 1 hypoglycemia	7 out of 351	2 depression (worsening), 1 flu syndrome disorder	2 out of 70
29060/448 (A), 2005	Paroxetine IR	1 myocardial infarction, 1 emotional lability	2 out of 104	1 uterine fibroids enlarged, 1 gall bladder disorder	2 out of 50
29060/448 (B), 2005	Paroxetine CR	3 emotional lability, 1 hepatocellular jaundice, 1 manic reaction	6 out of 102	1 dehydration, 1 accidental overdose	2 out of 51
29060/449 (A), 2005	Paroxetine IR	1 emotional lability, 1 abortion, 2 unintended pregnancy	3 out of 112 (4 SAE in 3 participants)	No serious adverse event	0 out of 55
29060/449 (B), 2005	Paroxetine CR	1 abdominal pain, 1 pancreatitis, 1 accidental over dose, 1 unintended pregnancy	2 out of 108 (4 SAE in 2 participants)	No serious adverse event	0 out of 55

Table 2 Summary of serious adverse events in the included trials *(Continued)*

PAR 29060.07.001, 2008	Paroxetine	1 acure depression, 1 acute alcohol intoxication and suicide ideation	2 out of 13	No serious adverse event	0 out of 12
Nyth et al., 1992	Citalopram	1 cerebral hemorrhage and death	1 out of 98	1 death	1 out of 51
Perahia et al., 2006	Paroxetine	1 back pain, 1 breast neoplasm	2 out of 97	No serious adverse events	0 out of 99
Mendels et al., 1999	Citalopram	1 prostatic hyper trophy, 1 bronchitis	2 out of 89	1 suicide	1 out of 91
NCT00668525 (A), 2010	Escitalopram	1 chest pain, 1 pharyngitis 1 multiple sclerosis	3 out of 319	1 asthma, 1 haemothorax	2 out of 108
NCT00668525 (B), 2010	Escitalopram	1 chest pain, 1 appendicitis, 2 anxiety, 1 suicidal ideation, 1 suicide attempt, 1 peripheral vasuclar disorder	7 out of 318	1 injury, 1 suicidal ideation	2 out of 107
Nierenberg et al., 2007	Escitalopram	1 death and 3 unspecified serious adverse events	4 out of 274	2 unspecified serious adverse events	2 out of 137
WELL AK130926, 2007	Escitalopram	1 agitation	1 out of 144	No serious adverse event	0 out of 132
WELL AK130927, 2007	Escitalopram	2 suicidal ideation, 1 hepatic function abnormal	3 out of 138	1 suicidal ideation, 1 sudden caridac death	2 out of 141
NCT01473381, 2014	Escitalopram	1 haemorrhagic anaemia, 1 diverticulitis, 1 ilium fracture, 1 road traffic accident, 1 traumatic renal injury, 1 wrist fracture, 1 abortion missed, 1 suicidal ideation, 1 hospitalisaation	6 out of 280 (9 SAE in 6 participants)	1 angina pectoris, 1 gastric disorder, 1 pneumonia, 1 neck abscess, 1 oral abscess, 1 abnormal electrocardiogram ST segment, 1 back pain, 1 suicidal ideation, 1 obstructive airways disorder	3 out of 281 (9 SAE in 3 participants)
Adamson et al., 2015	Citalopram	1 suicidal ideation, severe abdominal cramps	2 out of 73	No serious adverse event	0 out of 65
CIT-MD-03, 2005	Citalopram	2 congestive heart failure, 1 cerebro vascular accident, 1 hyponatremia	4 out of 84	1 cerebro vascular accident, 1 cellulitis	2 out of 90
29060/810 (A), 2005	Paroxetine CR 12.5 mg	2 abnormal laboratory value, 1 carcinoma of lung	3 out of 153	1 cerebro vascular disorder, 1 depression	1 out of 73 (2 SAE in 1 participants)
29060/810 (B), 2005	Paroxetine CR 25 mg	1 abnormal laboratory value, 1 gall bladder disorder, 1 anxiety, 1 emotional lability,	4 out of 148	1 pleura disorder, 1 sinusitis, 1 bronchitis	2 out of 73 (3 SAE in 2 participants)

SAE Serious adverse event

considered to exert effects on depressive symptoms. However, whether these effects are beneficial and clinically meaningful are the questions. Estimating a meaningful threshold for clinical significance is difficult and an assessment of clinical significance should ideally not only include a threshold on an assessment scale [182]. Major depressive disorder affects daily functioning, increases the risk of suicidal behaviour, and decreases quality of life [183]. Some adverse events might therefore be acceptable if SSRIs have clinically significant beneficial effects [13, 183, 184]. We therefore both predefined a threshold for clinical significance and assessed the balance between beneficial and harmful effects [13, 17, 184].

As threshold for clinical significance [14], we chose a drug-placebo difference of 3 points on the 17-item HDRS (ranging from 0 to 52 points) or an effect size of 0.50 standardised mean difference. This has been recommended by the National Institute for Clinical Excellence (NICE) in England and has been chosen in other reviews [4, 8, 31]. Nevertheless, these recommendations are not universally accepted and have been questioned [3]. Others have suggested the following 'rules of thumb' regarding the standardised mean difference: 0.2 a small effect, 0.5 a moderate effect, and 0.8 a large effect [16, 185]. One study has shown that a SSRI-placebo mean difference of up to three points on the HDRS corresponds to 'no clinical change' [186]. Another valid study has shown that a SSRI-placebo difference of 3 points is undetectable by clinicians, and that a mean difference of 7 HDRS points, or a standardized mean effect size of 0.875, is required to correspond to a rating of 'minimal improvement' [187]. It has been speculated that the 'placebo' response in antidepressant trials has been increasing during recent years [188]. If there is a

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 19 of 28

Table 3 Summary of all reported adverse events in the included trials

Event	No. of trials reporting the event	SSRI		Placebo		Relative risk (95% CI)	Number needed to harm (NNH)	P value
		Number of participants with the event	Number of participants randomised	Number of participants with the event	Number of participants randomised			
Abnormal ejaculation	15	183	3236	7	1903	5.43 [3.22, 9.14]	19	P < 0.00001
Tremor	28	301	3502	61	2929	3.16 [2.37, 4.21]	16	P < 0.00001
Anorexia	19	220	2350	42	1680	2.78 [2.03, 3.79]	15	P < 0.00001
Nausea	78	2524	12,257	779	8491	2.48 [2.22, 2.77]	9	P < 0.00001
Somnolence	59	1336	10,351	345	6674	2.25 [2.00, 2.53]	13	P < 0.00001
Sweating	34	440	5274	124	3478	2.20 [1.80, 2.70]	21	P < 0.00001
Asthenia	23	497	3968	155	2265	1.71 [1.43, 2.04]	18	P < 0.00001
Diarrhoea	58	1458	11,056	561	7099	1.66 [1.51, 1.83]	19	P < 0.00001
Constipation	50	606	6698	273	4892	1.60 [1.35, 1.89]	29	P < 0.00001
Insomnia	69	1500	11,934	582	7956	1.49 [1.35, 1.64]	19	P < 0.00001
Dizziness	55	849	8900	398	6161	1.39 [1.24, 1.57]	33	P < 0.00001
Dry mouth	73	1376	11,303	693	7904	1.37 [1.25, 1.49]	30	P < 0.00001
Libido decreased	8	78	1481	11	1083	3.48 [1.92, 6.32]	24	P < 0.0001
Sexual dysfunction	6	96	719	16	389	2.85 [1.77, 4.59]	11	P = 0.0001
Appetite decreased	8	68	932	24	885	2.63 [1.66, 4.17]	22	P < 0.0001
Fatigue	26	409	5098	153	3545	1.69 [1.32, 2.17]	27	P < 0.0001
Vomiting or upset stomach	20	189	2376	101	2037	1.55 [1.16, 2.08]	34	P = 0.003
Flu syndrome	7	57	1069	19	822	2.13 [1.28, 3.54]	34	P = 0.004
Drowsiness	5	38	253	19	256	1.90 [1.18, 3.04]	14	P = 0.004
Blurred/abnormal vision or dry eyes	17	116	1862	55	1566	1.55 [1.15, 2.10]	37	P = 0.004
Nervousness	22	484	3863	147	2043	1.35 [1.10, 1.66]	19	P = 0.004
Back pain	11	85	2404	71	1594	0.66 [0.48, 0.91]	109	P = 0.01
Headache	72	2386	11,085	1427	7805	1.08 [1.01, 1.14]	31	P = 0.02
Dyspepsia	23	331	4304	159	2956	1.29 [1.04, 1.59]	44	P = 0.02
Weight loss	3	26	562	9	560	2.48 [1.17, 5.25]	34	P = 0.02
Hypertension	4	17	933	25	761	0.51 [0.28, 0.93]	69	P = 0.03
Central or peripheral nervous system	4	104	221	16	57	1.58 [1.03, 2.43]	6	P = 0.04
Lightedness/faint feeling	3	9	147	1	144	4.81 [1.06, 21.72]	19	P = 0.04
Accidental injury	3	15	672	23	516	0.50 [0.25, 0.99]	45	P = 0.05

Table 3 Summary of all reported adverse events in the included trials (Continued)

Agitation	5	67	613	22	398	1.60 [1.01, 2.54]	19	$P = 0.05$
Impotence	3	19	868	1	603	3.12 [0.99, 9.88]	50	$P = 0.05$
Taste perversion	2	9	389	1	390	5.80 [1.02, 33.03]	49	$P = 0.05$
Shaking	2	7	121	0	116	7.19 [0.91, 57.03]	18	$P = 0.06$
Rhinitis	14	197	3004	171	1969	0.78 [0.63, 0.97]	47	$P = 0.06$
Palpitations	10	63	1572	30	1292	1.55 [0.97, 2.50]	60	$P = 0.07$
Infection	7	104	997	57	612	1.31 [0.97, 1.75]	90	$P = 0.08$
Amnesia	2	5	484	12	486	0.44 [0.16, 1.20]	70	$P = 0.11$
Psychiatric adverse effects	1	19	40	7	26	1.76 [0.87, 3.60]	5	$P = 0.12$
Sleep disturbance	3	15	240	7	223	1.98 [0.83, 4.73]	33	$P = 0.13$
Sinusitis	4	35	751	38	629	0.69 [0.43, 1.11]	73	$P = 0.13$
Urinary frequency	4	40	695	19	624	1.80 [0.82, 3.96]	37	$P = 0.13$
Anxiety	17	150	2983	89	2250	1.27 [0.91, 1.77]	94	$P = 0.15$
Appetie increased	3	5	335	10	330	0.49 [0.17, 1.43]	65	$P = 0.19$
Coughing	1	8	100	4	96	1.92 [0.60, 6.17]	27	$P = 0.27$
Tinnitus	2	4	319	1	319	2.84 [0.44, 18.44]	107	$P = 0.27$
Adverse events overall	4	32	232	14	160	1.36 [0.77, 2.39]	20	$P = 0.29$
Unpleasant taste	1	2	21	0	20	4.77 [0.24, 93.67]	11	$P = 0.30$
Congestive heart failure	1	0	335	2	336	0.20 [0.01, 4.16]	168	$P = 0.30$
Gastrointestial	4	66	194	14	91	1.53 [0.65, 3.62]	6	$P = 0.33$
Autonomic nervous system	1	10	40	4	26	1.63 [0.57, 4.64]	11	$P = 0.36$
Respiratory disorder	10	244	2764	107	1276	0.90 [0.73, 1.12]	227	$P = 0.36$
Vasodialtion	4	11	368	4	355	1.75 [0.44, 6.94]	54	$P = 0.43$
Flatulence	6	94	1763	34	972	1.24 [0.72, 2.16]	55	$P = 0.44$
Malaise	1	0	21	1	20	0.32 [0.01, 7.38]	20	$P = 0.48$
Depression aggrevated	1	6	337	1	180	1.95 [0.31, 12.35]	82	$P = 0.48$
Female genital disorders	1	7	310	1	149	1.93 [0.31, 12.10]	63	$P = 0.48$
Weight gain	1	3	129	5	129	0.60 [0.15, 2.46]	65	$P = 0.48$
Tachycardia	6	9	989	6	996	1.43 [0.52, 3.96]	326	$P = 0.49$
Arrhythmia	1	0	335	1	336	0.33 [0.01, 8.18]	336	$P = 0.50$
Atrial fibrillation	1	1	335	0	336	3.01 [0.12, 73.60]	335	$P = 0.50$
Abnormal electrocardiogram	1	1	335	0	336	3.01 [0.12, 73.60]	335	$P = 0.50$
Migraine	1	1	129	0	129	3.00 [0.12, 72.96]	129	$P = 0.50$
Chest discomfort	1	1	21	0	20	2.86 [0.12, 66.44]	21	$P = 0.51$
Rash	5	9	280	12	282	0.79 [0.37, 1.70]	97	$P = 0.55$
Vertigo	1	10	337	3	180	1.58 [0.32, 7.68]	77	$P = 0.57$
Dysuria	1	2	21	1	20	1.90 [0.19, 19.40]	23	$P = 0.58$
Pruritus (Itching)	3	8	187	10	185	0.79 [0.30, 2.10]	89	$P = 0.63$
Orthostatic hypotension	2	6	466	3	309	1.37 [0.37, 5.02]	316	$P = 0.64$
Upper respiratory tract infection	14	164	2380	123	1882	0.95 [0.75, 1.20]	282	$P = 0.68$
Body as a whole	1	6	40	3	26	1.30 [0.36, 4.75]	29	$P = 0.69$
Loose stools	1	3	100	2	96	1.44 [0.25, 8.43]	110	$P = 0.69$

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 21 of 28

Table 3 Summary of all reported adverse events in the included trials *(Continued)*

Gastritis	1	3	62	5	138	1.34 [0.33, 5.41]	83	$P = 0.69$
Cardiovascular disorder	2	10	444	3	355	0.81 [0.22, 2.95]	72	$P = 0.75$
Pain	5	65	893	59	729	0.95 [0.68, 1.34]	123	$P = 0.77$
Abnormal thinking	1	7	335	8	336	0.88 [0.32, 2.39]	344	$P = 0.80$
Abnormal acne	1	3	32	1	14	1.31 [0.15, 11.54]	45	$P = 0.81$
Confusion	4	5	640	5	621	0.87 [0.28, 2.74]	4187	$P = 0.81$
Myalgia	5	34	692	26	572	0.93 [0.45, 1.94]	272	$P = 0.85$
Irritability	6	18	643	20	642	0.92 [0.38, 2.27]	317	$P = 0.86$
Numbness	2	12	272	8	269	0.77 [0.04, 13.27]	70	$P = 0.86$
Abdominal pain	9	98	1967	52	1150	1.02 [0.73, 1.44]	218	$P = 0.89$
Trauma	2	47	883	14	267	0.95 [0.41, 2.20]	1261	$P = 0.90$
Eructation (burping)	1	5	100	5	96	0.96 [0.29, 3.21]	480	$P = 0.95$
Over sedation	3	24	677	7	514	1.91 [0.83, 4.39]	46	$P = 0.13$

'response' to placebo this has of course to be considered when interpreting a mean difference between drug and placebo. However, it is unlikely that depressed patient have a significant placebo effect [189] and it has recently been shown that the placebo response has been stable for 25 years [188]. Even based on our predefined minimal thresholds for clinical significance, the effects of SSRIs did not have a clinically meaningful effect on depressive symptoms. Furthermore, per our meta-analyses SSRIs significantly increase the risk of both serious and non-serious adverse events.

The best-worst and worst-best case scenarios showed that incomplete outcome data bias alone theoretically could have caused the apparent statistically significant beneficial effect of SSRIs. Furthermore, seen in the light of the total number of trials, only a relatively limited number of trials reported on each of our pre-defined outcomes. This increases the risk of selective outcome reporting bias. Apart from the high risk of incomplete outcome data bias and selective outcome reporting bias, all the included trials were assessed at high risk of bias. All trials used placebo as control intervention and due to the large number of adverse events, some patients might have figured out if they received an 'active' intervention or not, which might question the blinding of the trials. Nevertheless, it may be argued that our bias risk assessment often will lead to no trials with low risk of bias. However, similar bias risk assessments have been used in several previous systematic review (see, e.g., most Cochrane Hepato-Biliary Group systematic reviews) and our bias risk assessment is based on valid evidence clearly showing that if each of the used bias risk domains is 'high risk of bias' or 'unclear risk of bias' then there is a risk of overestimation of benefits and underestimation of harms [184, 190–197]. Furthermore, the risks of bias observed here just mirrors our experience in 786 randomised trials on depression [198].

We chose 'remission' as a primary outcome because we expected trialists to use this outcome frequently. To present a complete overview of the evidence on SSRIs for depression we also included 'no response' (less than 50% reduction on HDRS or MADRS during the intervention period) in a post hoc analysis because this outcome was frequently used in the included trials and by requests from peer reviewers. However, our results on no remission and no response should be interpreted with great caution for a number of reasons: 1) the assessments of remission and response were primarily based on single HDRS scores and it is questionable whether single HDRS scores are indications of full remission or adequate response to the intervention; 2) information is lost when continuous data are transformed to dichotomous data and the analysis results can be greatly influenced by the distribution of data and the choice of an arbitrary cut-point [16, 199–201]; 3) even though a larger proportion of participants cross the arbitrary cut-point in the SSRI group compared with the control group (often HDRS below 8 for remission and 50% HDRS reduction for response), the effect measured on HDRS might still be limited to a few HDRS points (e.g., 3 HDRS points) or less; 4) by only focusing on how many patients cross a certain line for benefit, investigators ignore how many patients are deteriorating at the same time. If results, e.g., show relatively large beneficial effects of SSRIs when remission and response are assessed but very small averaged effects (as our results show) – then it must be because similar proportions of the participants are harmed (increase on the HDRS compared to placebo) by SSRIs. Otherwise the averaged effect would not show small or no difference in effect. The clinical significance of our results on 'no remission' and 'no response' should therefore be questioned. The methodological limitations of using 'response' as an

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 22 of 28

outcome has been investigated in a valid study by Kirsch et al. who conclude that: "response rates based on continuous data do not add information, and they can create an illusion of clinical effectiveness" [202]. In retrospect, due to these methodological limitations we should not have assessed 'no remission' or 'no response' as outcomes. This is a clear limitation of our review [16, 199–201].

Our tests for subgroup difference comparing trials with a baseline HDRS score below and above 23 points and meta-regression showed that the effects of SSRIs seem to increase with increased baseline HDRS score. Others have also shown that trials randomising participants with a higher baseline HDRS mean average seem to show larger effects of antidepressants [7, 8]. However, it is difficult to interpret why trials with higher average baseline HDRS score seem to have a larger effect of SSRIs. This might just be due to random error. No matter, it cannot be concluded based on these results that SSRIs work better on more severely depressed patients. To make such a conclusion individual patient data would be necessary, i.e., it would be necessary to show that it is actually the patients with higher baseline HDRS scores who have the larger effects. Gibbons et al. used longitudinal person-level data from a large set of published and unpublished studies and showed baseline severity was not significantly related to degree of SSRI treatment advantage over placebo [3]. It must be noted that the intervention effects in the group with HDRS scores above 23 points were still below our threshold for clinical significance, supporting Gibbons and co-workers' results.

Leucht et al. have suggested that effects sizes of SSRIs in randomised clinical trials have declined over time [203]. Post-hoc meta-regression of the HDRS results confirmed their results (effect sizes going down from around 0.8 in the early 1980s to 0.25 in 2012). The reasons for the decreasing effect is not entirely understood but might be due to better methodology nowadays or recruitment of different types of participants [203]. Leucht et al. also suggested that a lack of difference between antidepressants and placebo is caused by an increasing 'placebo' effect (spontaneous recovery) [203]. This seem less important from a patient perspective, i.e., whether a certain drug should be used should be based on the benefits and harms of this drug compared with placebo. Furthermore, the increasing placebo effect has recently been severely questioned [188].

Our present systematic review has several strengths. Our protocol was registered prior to the systematic literature search in all relevant databases, data extraction, and data analyses [14]. Data were double-extracted by independent authors minimising the risk of inaccurate data extraction, and we assessed the risk of bias in all

trials according to Cochrane [16]. We used Trial Sequential Analysis to control the risks of random errors [25, 29, 204], and the analyses of the primary outcomes showed that the accrued information sizes were sufficient. Both visual assessments of forest plots and statistical test showed limited signs of statistical heterogeneity, e.g., I^2 was 0% when assessing risk of serious adverse events. Hence, these findings increase the validity of our review results and indicate that the effects shown are consistent across the different trials. Multiple previous reviews and meta-analyses have, as mentioned in our Background, assessed the effects of SSRIs and have generally concluded that SSRIs have significant effects on depressive symptoms [3–8]. However, the estimated results (and not the conclusions the review authors made) of these reviews and meta-analyses actually are in agreement with our present results and show that SSRIs do not seem to benefit patients more than a few HDRS points. This increases the validity of our present results. Furthermore, we assessed in detail the risks of serious adverse events and of non-serious adverse events and found that both were significantly increased by SSRIs.

Our systematic review has several limitations. Our HDRS mean differences were averaged effects. Hence, it cannot be concluded that SSRIs do not have clinically significant effects on all depressed participant. E.g., certain severely depressed patients compared with lightly depressed patients (e.g., so-called professional patients or symptomatic volunteers [203]) might benefit from SSRIs even though there is no evidence backing this hypothesis. However, any clinical research result will have this 'limitation'. Specific patients might benefit from any given intervention even though valid research results have shown that this intervention 'on average' is ineffective or even harmful. All trials were at high risk of bias per several bias risk domains and especially the risk of incomplete outcome data, selective outcome reporting, and insufficient blinding bias may bias our review results. Our GRADE assessments show that due to the high risks of bias the quality of the evidence must be regarded as very low. The high risks of bias question the validity of our meta-analysis results as high risk of bias trials tend to overestimate benefits and underestimate harms [194, 205]. The 'true' effect of SSRIs might not even be statistically significant.

We chose to include all SSRIs in our primary analysis. We did this to increase the statistical power and precision and to be able to compare the effects of the different SSRIs in subgroup analysis. Comparing the different SSRIs in test for subgroup differences did not show significant differences, indicating the effects (or lack of effects) of the different SSRIs are similar. Nevertheless, we cannot rule out that certain SSRIs may have beneficial or harmful effects that we have not identified in this

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 23 of 28

review due to lack of relevant data. We identified very limited data on the effects of SSRIs on long-term outcomes, suicidal behaviour, and quality of life, so the effects of SSRIs on these outcomes are unclear. E.g., we only identified six trials assessing quality of life which substantially increase the risk of selective outcome reporting bias and thereby limit the validity of the meta-analysis result. Furthermore, the trialists did not use the same questionnaire. Quality of life is without question an outcome with great relevance to the patient and we urge future trialists to assess quality of life. However, any given quality of life questionnaire must be validated (shown to be correlated to, e.g., suicidal behaviour or other clinical events) before valid conclusions may be drawn based on this outcome. It must be shown that scores on a given questionnaire do reflect the actual 'quality of life'. Valid consensus on choosing the optimal quality of life assessment method does not exist and this is a limitation of assessing quality of life in depressed patients. Our eight-step procedure used to assess if the thresholds for statistical and clinical significance are crossed, is based on generally accepted and validated methodology but the use of the eight-step procedure has not yet been validated in simulation studies or empirical studies [12, 13]. Even though the eight-step procedure has been used in several systematic reviews it is not universally accepted. This may be a limitation of our methodology.

The Committee for Medicinal Products for Human Use (CHMP) concluded"......... that, as no public health concerns have been identified, no regulatory action is necessary on the basis of Kirsch et al.'s findings" when the latter team questioned the benefits of antidepressants [182]. Per our results, we now believe that there is valid evidence for a public concern regarding the effects of SSRIs. We agree with Andrews et al. that that antidepressants seem to do more harm than good [206]. We have clearly shown that SSRIs significantly increase the risks of both serious and several non-serious adverse events. The observed harmful effects seem to outweigh the potential small beneficial clinical effects of SSRIs, if they exist. Our results confirm the findings from other studies questioning the effects of SSRIs [8, 207], but are in contrast to the results of other reviews concluding that SSRIs are effective interventions for depression [3, 6, 10, 208]. However, our present analyses represent the most comprehensive systematic review on the topic and we hope it may guide clinical practice.

Conclusions

SSRIs versus placebo seem to have statistically significant effects on depressive symptoms, but the clinical significance of these effects seems questionable and all trials were at high risk of bias. Furthermore, SSRIs versus placebo significantly increase the risk of both serious and non-serious adverse events. Our results show that the harmful effects of SSRIs versus placebo for major depressive disorder seem to outweigh any potentially small beneficial effects.

Additional files

Additional file 1: Search strategies. (DOC 41 kb)

Additional file 2: List of excluded trials. (TXT 2726 kb)

Additional file 3: Table S1. Trial characteristics. (XLS 80 kb)

Additional file 4: Figure S1. Subgroup analysis of for profit bias. (PNG 56 kb)

Additional file 5: Summary of findings table. (DOCX 72 kb)

Additional file 6: Summary of the 20 Most Common Adverse Events in the Included Trials. (PDF 37 kb)

Additional file 7: Figure S3. Meta-analysis of abnormal ejaculation. (PDF 203 kb)

Additional file 8: Figure S4. Meta-analysis of tremor. (PDF 240 kb)

Additional file 9: Figure S5. Meta-analysis of anorexia. (PDF 194 kb)

Additional file 10: Figure S6. Meta-analysis of nausea. (PDF 579 kb)

Additional file 11: Figure S7. Meta-analysis of somnolence. (PDF 459 kb)

Additional file 12: Figure S8. Meta-analysis of sweating. (PDF 283 kb)

Additional file 13: Figure S9. Meta-analysis of asthenia. (PDF 234 kb)

Additional file 14: Figure S10. Meta-analysis of diarrhoea. (PDF 475 kb)

Additional file 15: Figure S11. Meta-analysis of constipation. (PDF 369 kb)

Additional file 16: Figure S12. Meta-analysis of insomnia. (PDF 538 kb)

Additional file 17: Figure S13. Meta-analysis of dizziness. (PDF 406 kb)

Additional file 18 Figure S14. Meta-analysis of dry mouth. (PDF 542 kb)

Additional file 19: Figure S15. Meta-analysis of libido decreased. (PDF 136 kb)

Additional file 20: Figure S16. Meta-analysis of sexual dysfunction. (PDF 124 kb)

Additional file 21: Figure S17. Meta-analysis of appetite decreased. (PDF 123 kb)

Additional file 22: Figure S18. Meta-analysis of fatigue. (PDF 239 kb)

Additional file 23: Figure S19. Meta-analysis of vomiting or stomach upset. (PDF 190 kb)

Additional file 24: Figure S20. Meta-analysis of flu syndrome. (PDF 130 kb)

Additional file 25: Figure S21. Meta-analysis of drowsiness. (PDF 104 kb)

Additional file 26: Figure S22. Meta-analysis of blurred or abnormal vision. (PDF 166 kb)

Additional file 27: Figure S23. Meta-analysis of nervousness. (PDF 224 kb)

Additional file 28: Figure S24. Meta-analysis of back pain. (PDF 153 kb)

Additional file 29: Figure S25. Meta-analysis of headache. (PDF 226 kb)

Additional file 30: Figure S26 Meta-analysis of dyspepsia. (PDF 226 kb)

Additional file 31: Figure S27 Meta-analysis of weight loss. (PDF 97 kb)

Additional file 32: Figure S2 Meta-analysis of no response. (PNG 105 kb)

Jakobsen et al. BMC Psychiatry (2017) 17:58

Page 24 of 28

Availability of data and materials
We have included all data in the review publication and in the additional files, but if additional information is needed Janus Christian Jakobsen can always be contacted (jcj@ctu.dk).

Authors' contributions
JCJ wrote the first draft and had full access to all the data in the review and takes responsibility for the integrity of the data and the accuracy of the data analysis. Selection of trials and data extraction: KKK, AS, SGH, SES, KLM, MI, MBB, IJP, JK, SLK, SE, AT, and JCJ. Disagreements were resolved by discussion or through arbitration by JCJ. Statistical analyses: JCJ, KKK, and JL. JCJ, JL and CG contributed to the development of the protocol and the review. All authors contributed to and have approved the manuscript.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
Not applicable.

Author details
[1]The Copenhagen Trial Unit, Centre for Clinical Intervention Research, Department 7812 Rigshospitalet, Copenhagen University Hospital, Blegdamsvej 9, Rigshospitalet, DK 2100 Copenhagen, Denmark. [2]Department of Cardiology, Holbæk Hospital, Holbæk, Denmark. [3]Mental Health Centre Copenhagen, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark.

References
1. Gualano MR, Bert F, Mannocci A, La Torre G, Zeppegno P, Siliquini R. Consumption of antidepressants in Italy: recent trends and their significance for public health. Psychiatr Serv. 2014;65(10):1226–31.
2. Wise J. GPs in England prescribed 2.7 million extra antidepressants during 2012 recession. BMJ. 2014;348:g3607.
3. Gibbons RD, Hur K, Brown CH, Davis JM, Mann JJ. Benefits from antidepressants: synthesis of 6-week patient-level outcomes from double-blind placebo-controlled randomized trials of fluoxetine and venlafaxine. Arch Gen Psychiatry. 2012;69(6):572–9.
4. Fournier JC, DeRubeis RJ, Hollon SD, Dimidjian S, Amsterdam JD, Shelton RC, Fawcett J. Antidepressant drug effects and depression severity: a patient-level meta-analysis. JAMA. 2010;303(1):47–53.
5. Khan A, Leventhal RM, Khan SR, Brown WA. Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration database. J Clin Psychopharmacol. 2002;22(1):40–5.
6. Undurraga J, Baldessarini RJ. Randomized, placebo-controlled trials of antidepressants for acute major depression: thirty-year meta-analytic review. Neuropsychopharmacol. 2012;37(4):851–64.
7. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. N Engl J Med. 2008;358(3):252–60.
8. Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. PLoS Med. 2008;5(2):e45.
9. Arroll B, Elley CR, Fishman T, Goodyear-Smith FA, Kenealy T, Blashki G, Kerse N, MacGillivray S. Antidepressants versus placebo for depression in primary care. Cochrane Database of Syst Rev. 2009.
10. Wilson K, Mottram PG, Sivananthan A, Nightingale A. Antidepressants versus placebo for the depressed elderly. Cochrane Database Syst Rev. 2001.
11. Englev E, Petersen KP. ICH-GCP Guideline: quality assurance of clinical trials. Status and perspectives. Dan Med J. 2003;165(16):1659–62.
12. Jakobsen JC, Gluud C, Winkel P, Lange T, Wetterslev J. The thresholds for statistical and clinical significance - a five-step procedure for evaluation of intervention effects in randomised clinical trials. BMC Med Res Methodol. 2014;14:34.
13. Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. BMC Med Res Methodol. 2014;14(1):120.
14. Jakobsen JC, Lindschou J, Hellmuth S, Schou A, Krogh J, Gluud C. The effects of selective serotonin reuptake inhibitors versus no intervention, placebo, or 'active' placebo in patients with major depressive disorder. A systematic review of randomised clinical trials with meta-analyses and trial sequential analyses. PROSPERO 2013:CRD42013004420 Available from http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42013004420.
15. Moncrieff J, Wessely S, Hardy R. Active placebos versus antidepressants for depression. Coch Database Syst Rev. 2004;1:CD003012.
16. Higgins JPT, Green S. The Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0. The Cochrane Collaboration 2011, Available from http://www.cochrane-handbook.org/.
17. Jakobsen JC, Lindschou J, Hellmuth S, Schou A, Krogh J, Gluud C. The effects of selective serotonin reuptake inhibitors versus no intervention, placebo, or 'active' placebo in patients with major depressive disorder. A systematic review of randomised clinical trials with meta-analyses and trial sequential analyses. PROSPERO 2013:CRD42013004420 Available from http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42013004420 2013.
18. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry. 1960;23(3):56–61.
19. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry. 1979;134:382–9.
20. Beck A, Steer RA, Brown GK. Beck depression inventory—second edition: manual. San Antonio: Psychological Corporation. 1996;4:561–71.
21. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64(4):383–94.
22. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177–88.
23. The Nordic Cochrane Centre (The Cochrane Collaboration): Review Manager (RevMan) [Computer program]. Version 5.3. In.; 2008.
24. Brok J, Thorlund K, Gluud C, Wetterslev J. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analysis. J Clin Epidemiol. 2008;61:763–9.
25. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. J Clin Epidemiol. 2008;61(1):64–75.
26. Thorlund K, Devereaux PJ, Wetterslev J, Guyatt G, Ioannidis JP, Thabane L, Gluud LL, Als-Nielsen B, Gluud C. Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses? Int J Epidemiol. 2009;38(1):276–86.
27. Thorlund K, Imberger G, Walsh M, Chu R, Gluud C, Wetterslev J, Guyatt G, Devereaux PJ, Thabane L. The number of patients and events required to limit the risk of overestimation of intervention effects in meta-analysis – a simulation study. PLoS One. 2011;6:e25491.
28. Thorlund K, Engstrøm J, Wetterslev J, Brok J, Imberger G, Gluud C. User manual for trial sequential analysis (TSA). Copenhagen: Copenhagen Trial Unit, Centre for Clinical Intervention Research; 2011. p. 1–115. Available from http://www.ctu.dk/tsa.
29. Wetterslev J, Thorlund K, Brok J, Gluud C. Estimating required information size by quantifying diversity in random-effects model meta-analyses. BMC Med Res Methodol. 2009;9:86.
30. Higgins J. Convincing evidence from controlled and uncontrolled studies on the lipid-lowering effect of a statin [editorial]. Cochrane Database Syst Rev. 2012;12:ED000049.
31. National Institute for Clinical Excellence. Depression: Mangement of Depression in Primary and Secundary Care. London: National Institute for Clinical Excellence; 2004. p. 640.

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 25 of 28

32. Byerley WF, Reimherr FW, Wood DR, Grosser BI. Fluoxetine, a selective serotonin uptake inhibitor, for the treatment of outpatients with major depression. J Clin Psychopharmacol. 1988;8:112–5.

33. Davidson JRT, Gadde KM, Fairbank JA, Ranga Rama KK, Califf RM, Binanay C, Parker CB, Pugh N, Hartwell TD, Vitiello B, et al. Effect of Hypericum perforatum (St John's wort) in major depressive disorder: A randomized controlled trial. JAMA. 2002;287(14):10.

34. Fava M, Alpert J, Nierenberg AA, Mischoulon D, Otto MW, Zajecka J, Murck H, Rosenbaum JF. A Double-blind, randomized trial of St John's wort, fluoxetine, and placebo in major depressive disorder. J Clin Psychopharmacol. 2005;25:441–7.

35. Fava M, Amsterdam JD, Deltito JA, Salzman C, Schwaller M, Dunner DL. A double-blind study of paroxetine, fluoxetine, and placebo in outpatients with major depression. Ann Clin Psychiatry. 1998;10:145–50.

36. Feighner JP, Boyer WF, Merideth CH, Hendrickson GG. A double-blind comparison of fluoxetine, imipramine and placebo in outpatients with major depression. Int Clin Psychopharmacol. 1989;4:127–34.

37. Godlewska BR, Norbury R, Selvaraj S, Cowen PJ, Harmer CJ. Short-term SSRI treatment normalises amygdala hyperactivity in depressed patients. Psychol Med. 2012;42:2609–17.

38. Hunter AM, Cook IA, Greenwald SD, Tran ML, Miyamoto KN, Leuchter AF. The antidepressant treatment response index and treatment outcomes in a placebo-controlled trial of fluoxetine. J Clin Neurophysiol. 2011;28:478–82.

39. Itil TM, Shrivastava RK, Mukherjee S. A double-blind placebo-controlled study of fluvoxamine and imipramine in out-patients with primary depression. Brit J Clin Pharmacol. 1983;15(Suppl#3):1983.

40. Jindal RD, Friedman ES, Berman SR, Fasiczka AL, Howland RH, Thase ME. Effects of sertraline on sleep architecture in patients with depression. J Clin Psychopharmacol. 2003;23:540–8.

41. Katz MM, Tekell JL, Bowden CL, Brannan S, Houston JP, Berman N, Frazer A. Onset and early behavioral effects of pharmacologically different antidepressants and placebo in depression. Neuropsychopharmacol. 2004; 29:566–79.

42. McGrath PJ, Stewart JW, Janal MN, Petkova E, Quitkin FM, Klein DF. A placebo-controlled study of fluoxetine versus imipramine in the acute treatment of atypical depression. Am J Psychiatry. 2000;157:344–50.

43. Muijen M, Roy D, Silverstone T, Mehmet A, Christie M. A comparative clinical trial of fluoxetine, mianserin and placebo in depressed outpatients. Acta Psychiatr Scand. 1988;78:384–90.

44. Petrakis I, Carroll KM, Nich C, Gordon L, Kosten T, Rounsaville B. Fluoxetine treatment of depressive disorders in methadone-maintained opioid addicts. Drug Alcohol Depend. 1998;50:221–6.

45. Rapaport MH, Schneider LS, Dunner DL, Davies JT, Pitts CD. Efficacy of controlled-release paroxetine in the treatment of late-life depression. J Clin Psychiatry. 2003;64:1065–74.

46. Roth D, Mattes J, Sheehan KH, Sheehan DV. A double-blind comparison of fluvoxamine, desipramine and placebo in outpatients with depression. Prog Neuro Psychopharmacol Biol Psychiat. 1990;14:929–39.

47. Roy A. Placebo-controlled study of sertraline in depressed recently abstinent alcoholics. Biol Psychiatry. 1998;44:633–7.

48. Schneider LS, Nelson JC, Clary CM, Newhouse P, Krishnan KR, Shiovitz T, Weihs K, Sertraline Elderly Depression Study G. An 8-week multicenter, parallel-group, double-blind, placebo-controlled study of sertraline in elderly outpatients with major depression. Am J Psychiatry. 2003;160:1277–85.

49. Sheehan DV, Nemeroff CB, Thase ME, Entsuah R: Placebo-controlled inpatient comparison of venlafaxine and fluoxetine for the treatment of major depression with melancholic features. Int Clin Psychopharmacol 2009, 24(2):March.

50. Mischoulon D, Price LH, Carpenter LL, Tyrka AR, Papakostas GI, Baer L, Dording CM, Clain AJ, Durham K, Walker R, et al. A double-blind, randomized, placebo-controlled clinical trial of S-adenosyl-L-methionine (SAMe) versus escitalopram in major depressive disorder. J Clin Psychiatry. 2014;75(4):370–6.

51. Barber JP, Barrett MS, Gallop R, Rynn MA, Rickels K. Short-term dynamic psychotherapy versus pharmacotherapy for major depressive disorder: a randomized, placebo-controlled trial. J Clin Psychol. 2012;73(1):66–73.

52. Binneman B, Feltner D, Kolluri S, Shi Y, Qiu R, Stiger T. A 6-week randomized, placebo-controlled trial of CP-316,311 (a selective CRH1 antagonist) in the treatment of major depression. Am J Psychiatry. 2008;165: 617–20.

53. Blumenthal JA, Babyak MA, Doraiswamy PM, Watkins L, Hoffman BM, Barbour KA, Herman S, Craighead WE, Brosse AL, Waugh R, et al. Exercise and pharmacotherapy in the treatment of major depressive disorder. Psychosomatic Med. 2007;69:587–96.

54. Bose A, Li D, Gandhi C. Escitalopram in the acute treatment of depressed patients aged 60 years or older. Am J Geriatr Psychiatry. 2008;16:14–20.

55. Detke MJ, Wiltse CG, Mallinckrodt CH, McNamara RK, Demitrack MA, Bitter I. Duloxetine in the acute and long-term treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. Eur Neuropsychopharmacol. 2004;14(6):457–70.

56. Dube S, Dellva MA, Jones M, Kielbasa W, Padich R, Saha A, Rao P: A study of the effects of LY2216684, a selective norepinephrine reuptake inhibitor, in the treatment of major depression. J Psychiatric Res 2010, 44(6):April.

57. Edwards JG, Goldie A. Placebo-controlled trial of paroxetine in depressive illness. Hum Psychopharmacol. 1993;8(3):1993.

58. Fabre LF. A 6-week, double-blind trial of paroxetine, imipramine, and placebo in depressed outpatients. J Clin Psychiatry. 1992;53(Suppl):40–3.

59. Fabre LF, Abuzzahab FS, Amin M, Claghorn JL, Mendels J, Petrie WM, Dub S, Small JG. Sertraline safety and efficacy in major depression: A double-blind fixed-dose comparison with placebo. Biol Psychiatry. 1995; 38(9):1995.

60. Higuchi T, Hong JP, Jung HY, Watanabe Y, Kunitomi T, Kamijima K. Paroxetine controlled-release formulation in the treatment of major depressive disorder: a randomized, double-blind, placebo-controlled study in Japan and Korea. Psychiatry Clin Neurosci. 2011;65(7):655–63.

61. Kranzler HR, Mueller T, Cornelius J, Pettinati HM, Moak D, Martin PR, Anthenelli R, Brower KJ, O'Malley S, Mason BJ, et al. Sertraline treatment of co-occurring alcohol dependence and major depression. J Clin Psychopharmacol. 2006;26:13–20.

62. Lydiard RB, Stahl SM, Hertzman M, Harrison WM. A double-blind, placebo-controlled study comparing the effects of sertraline versus amitriptyline in the treatment of major depression. J Clin Psychiatry. 1997;58:484–91.

63. Rapaport MH, Lydiard RB, Pitts CD, Schaefer D, Bartolic EI, Iyengar M, Carfagno M, Lipschitz A. Low doses of controlled-release paroxetine in the treatment of late-life depression: a randomized, placebo-controlled trial. J Clin Psychiatry. 2009;70:46–57.

64. Reimherr FW, Chouinard G, Cohn CK, Cole JO, Itil TM, Lapierre YD, Masco HL, Mendels J. Antidepressant efficacy of sertraline: a double-blind, placebo- and amitriptyline-controlled, multicenter comparison study in outpatients with major depression. J Clin Psychiatry. 1990;51(Suppl B):18–27.

65. Sramek JJ, Kashkin K, Jasinsky O, Kardatzke D, Kennedy S, Cutler NR. Placebo-controlled study of ABT-200 versus fluoxetine in the treatment of major depressive disorder. Depress. 1995;3(4):1995.

66. Burke WJ, Gergel I, Bose A. Fixed-dose trial of the single isomer SSRI escitalopram in depressed outpatients. J Clin Psychiatry. 2002;63:331–6.

67. Cassano GB, Conti L, Massimetti G, Mengali F, Waekelin JS, Levine J. Use of a standardized documentation system (BLIPS/BDP) in the conduct of a multicenter international trial comparing fluvoxamine, imipramine, and placebo. Psychopharmacol Bull. 1986;22:52–8.

68. Cohn JB, Wilcox C. A comparison of fluoxetine, imipramine, and placebo in patients with major depressive disorder. J Clin Psychiatry. 1985;46:26–31.

69. Coleman CC, King BR, Bolden WC, Book MJ, Segraves RT, Richard N, Ascher J, Batey S, Jamerson B, Metz A. A placebo-controlled comparison of the effects on sexual functioning of bupropion sustained release and fluoxetine. Clin Therapeutics. 2001;23:1040–58.

70. Corrigan MH, Denahan AQ, Wright CE, Ragual RJ, Evans DL. Comparison of pramipexole, fluoxetine, and placebo in patients with major depression. Depress Anxiety. 2000;11:58–65.

71. DeRubeis RJ, Hollon SD, Amsterdam JD, Shelton RC, Young PR, Salomon RM, O'Reardon JP, Lovett ML, Gladis MMBL, Gallop R. Cognitive therapy vs medications in the treatment of moderate to severe depression. Arch Gen Psychiatry. 2005;62:409–16.

72. Dominguez RA, Goldstein BJ, Jacobson AF, Steinbook RM. A double-blind placebo-controlled study of fluvoxamine and imipramine in depression. J Clin Psychiatry. 1985;46:84–7.

73. Fabre LF, Crismon L. Efficacy of fluoxetine in outpatients with major depression. Curr Ther Res Clin Exp. 1985;37(1):1985.

74. Lydiard RB, Laird LK, Morton Jr, Steele TE, Kellner C, Laraia MT, Ballenger JC. Fluvoxamine, imipramine, and placebo in the treatment of depressed outpatients: Effects on depression. Psychopharmacol Bull. 1989;25(1):68.

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 26 of 28

75. March JS, Kobak KA, Jefferson JW, Mazza J, Greist JH. A double-blind, placebo-controlled trial of fluvoxamine versus imipramine in outpatients with major depression. J Clin Psychiatry. 1990;51:200–2.

76. Mayberg HS, Silva JA, Brannan SK, Tekell JL, Mahurin RK, McGinnis S, Jerabek PA. The functional neuroanatomy of the placebo effect. Am J Psychiatry. 2002;159:728–37.

77. Miller SM, Naylor GJ, Murtagh M, Winslow G. A double-blind comparison of paroxetine and placebo in the treatment of depressed patients in a psychiatric outpatient clinic. Acta Psychiatr Scand. 1989;350:143–4.

78. Nemeroff CB, Thase ME: A double-blind, placebo-controlled comparison of venlafaxine and fluoxetine treatment in depressed outpatients. J Psychiatr Res 2007, 41(3–4):April/June.

79. Norton KR, Sireling LI, Bhat AV, Rao B, Paykel ES: A double-blind comparison of fluvoxamine, imipramine and placebo in depressed patients. J Affect Disord 1984, 7(3–4):Dec.

80. Rickels K, Amsterdam JD, Avallone MF. Fluoxetine in major depression: A controlled study. Curr Ther Res Clin Exp. 1986;39(4):1986.

81. Rudolph RL, Feiger AD. A double-blind, randomized, placebo-controlled trial of once-daily venlafaxine extended release (XR) and fluoxetine for the treatment of depression. J Affect Disord. 1999;56:171–81.

82. Schatzberg A, Roose S. A double-blind, placebo-controlled study of venlafaxine and fluoxetine in geriatric outpatients with major depression. Am J Geriatr Psychiatry. 2006;14:361–70.

83. Shrivastava RK, Shrivastava SH, Overweg N, Blumhardt CL. A double-blind comparison of paroxetine, imipramine, and placebo in major depression. J Clin Psychiatry. 1992;53(Suppl):48–51.

84. Smith WT, Glaudin V. A placebo-controlled trial of paroxetine in the treatment of major depression. J Clin Psychiatry. 1992;53(Suppl):36–9.

85. Tyrer P, Marsden CA, Casey P, Seivewright N. Clinical efficacy of paroxetine in resistant depression. J Psychopharmacol. 1987;1(4):1987.

86. Fabre LF, Putman HP. A fixed-dose clinical trial of fluoxetine in outpatients with major depression. J Clin Psychiatry. 1987;48:406–8.

87. Feighner JP, Boyer WF. Paroxetine in the treatment of depression: a comparison with imipramine and placebo. Acta Psychiatr Scand. 1989;350:125–9.

88. Feighner JP, Boyer WF, Meredith CH, Hendrickson GG. A placebo-controlled inpatient comparison of fluvoxamine maleate and imipramine in major depression. Int Clin Psychopharmacol. 1989;4:239–44.

89. Claghorn JL. The safety and efficacy of paroxetine compared with placebo in a double-blind trial of depressed outpatients. J Clin Psychiatry. 1992;53(Suppl):33–5.

90. Claghorn JL, Earl CQ, Walczak DD, Stoner KA, Wong LF, Kanter D, Houser VP. Fluvoxamine maleate in the treatment of depression: a single-center, double-blind, placebo-controlled comparison with imipramine in outpatients. J Clin Psychopharmacol. 1996;16:113–20.

91. Cohn JB, Crowder JE, Wilcox CS, Ryan PJ. A placebo- and imipramine-controlled study of paroxetine. Psychopharmacol Bull. 1990;26:185–9.

92. Dunbar GC, Cohn JB, Fabre LF, Feighner JP, Fieve RR, Mendels J, Shrivastava RK. A comparison of paroxetine, imipramine and placebo in depressed out-patients. Brit J Psychiatry. 1991;159:394–8.

93. Fabre L, Birkhimer LJ, Zaborny BA, Wong LF, Kapik BM. Fluvoxamine versus imipramine and placebo: a double-blind comparison in depressed patients. Int Clin Psychopharmacol. 1996;11:119–27.

94. Feighner JP, Overö K. Multicenter, placebo-controlled, fixed-dose study of citalopram in moderate-to-severe depression. J Clin Psychiatry. 1999;60:824–30.

95. Griebel G, Beeské S, Stahl SM. The vasopressin V(1b) receptor antagonist SSR149415 in the treatment of major depressive and generalized anxiety disorders: results from 4 randomized, double-blind, placebo-controlled studies. J Clin Psychiatry. 2012;73:1403–11.

96. Kiev A. A double-blind, placebo-controlled study of paroxetine in depressed outpatients. J Clin Psychiatry. 1992;53(Suppl):27–9.

97. Olie JP, Gunn KP, Katz E. A double-blind placebo-controlled multicentre study of sertraline in the acute and continuation treatment of major depression. Eu Psychiatry. 1997;12(1):1997.

98. Peselow ED, Stanley M, Filippi AM, Barouche F, Goodnick P, Fieve RR. The predictive value of the dexamethasone suppression test. A placebo-controlled study. Brit J Psychiatry. 1989;155:667–72.

99. Rickels K, Amsterdam J, Clary C, Fox I, Schweizer E, Weise C. The efficacy and safety of paroxetine compared with placebo in outpatients with major depression. J Clin Psychiatry. 1992;53(Suppl):30–2.

100. Stahl SM. Placebo-controlled comparison of the selective serotonin reuptake inhibitors citalopram and sertraline. Biol Psychiatry. 2000;48:894–901.

101. Trivedi MH, Pigotti TA, Perera P, Dillingham KE, Carfagno ML, Pitts CD. Effectiveness of low doses of paroxetine controlled release in the treatment of major depressive disorder. J Clin Psychiatry. 2004;65:1356–64.

102. Forest Laboratories Inc: SCT-MD-01 (Unpublished Data) Forest Laboratories Inc. Study report from U.S. Food and Drug Administration. 2001.

103. Hoffman BM, Blumenthal JA, Babyak MA, Smith PJ, Rogers SD, Doraiswamy PM, Sherwood A. Exercise fails to improve neurocognition in depressed middle-aged and older adults. Med Sci Sports Exerc. 2008;40:1344–52.

104. Brunoni AR, Valiengo L, Baccaro A, Zanão TA, de Oliveira JF, Goulart A, Boggio PS, Lotufo PA, Benseñor IM, Fregni F. The sertraline vs. electrical current therapy for treating depression clinical study: results from a factorial, randomized, controlled trial. JAMA Psychiat. 2013;70:383–91.

105. Kasper S, Ebert B, Larsen K, Tonnoir B. Combining escitalopram with gaboxadol provides no additional benefit in the treatment of patients with severe major depressive disorder. European Neuropsychopharmacol (conference Publication). 2011;21:S408.

106. Wade A, Michael LO, Bang HK. Escitalopram 10 mg/day is effective and well tolerated in a placebo-controlled study in depression in primary care. Int Clin Psychopharmacol. 2002;17:95–102.

107. Silverstone PH, Ravindran A. Once-daily venlafaxine extended release (XR) compared with fluoxetine in outpatients with depression and anxiety. Venlafaxine XR 360 Study Group. J Clin Psychiatry. 1999;60:22–8.

108. Andreoli V, Caillard V, Deo RS, Rybakowski JK, Versiani M. Reboxetine, a new noradrenaline selective antidepressant, is at least as effective as fluoxetine in the treatment of depression. J Clin Psychopharmacol. 2002;22:393–9.

109. O'Flynn K, O'Keane V, Lucey JV, Dinan TG. Effect of fluoxetine on noradrenergic mediated growth hormone release: a double blind, placebo-controlled study. Biol Psychiatry. 1991;30:377–82.

110. Reynolds CF, Smith GS, Dew MA, Mulsant BH, Miller MD, Schlernitzauer M, Stack JA, Houck PR, Pollock BG. Accelerating symptom-reduction in late-life depression: a double-blind, randomized, placebo-controlled trial of sleep deprivation. Am J Geriatr Psychiatry. 2005;13:353–8.

111. Roose SP, Sackeim HA, Krishnan KR, Pollock BG, Alexopoulos G, Lavretsky H, Katz IR, Hakkarainen H, Old Old Depression Study G. Antidepressant pharmacotherapy in the treatment of depression in the very old: a randomized, placebo-controlled trial. Am J Psychiatry. 2004;161:2050–9.

112. Wang G, McIntyre A, Eriksson H. A randomized, double-blind study of the efficacy and tolerability of extended-release quetiapine fumarate (quetiapine XR) monotherapy in patients with major depressive disorder. Neuropsychiatr Dis Treat. 2014, 10.

113. Kasper S, de Swart H, Friis AH. Escitalopram in the treatment of depressed elderly patients. Am J Geriatr Psychiatry. 2005;13:884–91.

114. Learned S, Graff O, Roychowdhury S, Moate R, Krishnan KR, Archer G, Modell JG, Alexander R, Zamuner S, Lavergne A, et al. Efficacy, safety, and tolerability of a triple reuptake inhibitor GSK372475 in the treatment of patients with major depressive disorder: two randomized, placebo- and active-controlled clinical trials. J Psychopharmacol. 2012;26:653–62.

115. Mancino MJ, McGaugh J, Chopra MP, Guise JB, Cargile C, Williams DK, Thostenson J, Kosten TR, Sanders N, Oliveto A. Clinical efficacy of sertraline alone and augmented with gabapentin in recently abstinent cocaine-dependent patients with depressive symptoms. J Clin Psychopharmacol. 2014;34:234–9.

116. Pettinati HM, Oslin DW, Kampman KM, Dundon WD, Xie H, Gallis TL, Dackis CA, O'Brien CP. A double-blind, placebo-controlled trial combining sertraline and naltrexone for treating co-occurring depression and alcohol dependence. Am J Psychiatry. 2010;167:668–75.

117. Ratti E, Bellew K, Bettica P, Bryson H, Zamuner S, Archer G, Squassante L, Bye A, Trist D, Krishnan KR, et al. Results from 2 randomized, double-blind, placebo-controlled studies of the novel NK1 receptor antagonist casopitant in patients with major depressive disorder. J Clin Psychopharmacol. 2011;31:727–33.

118. Ravindran AV, Teehan MD, Bakish D, Yatham L, Oreilly R, Fernando ML, Manchanda R, Charbonneau Y, Buttars J. The impact of sertraline, desipramine, and placebo on psychomotor functioning in depression. Human Psychopharmacol. 1995;10(4):273–81.

119. 315 (Unpublished Data). Study report from U.S. Food and Drug Administration. In.; 1991.

120. Adamson SJ, Sellman JD, Foulds J, Frampton CM, Deering D, Dunn A, Berks J, Nixon L, Cape G. A randomized trial of combined citalopram and naltrexone for nonabstinent outpatients with co-occurring alcohol dependence and major depression. J Clin Psychopharmacol. 2015;35:143–9.

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 27 of 28

121. Ball WA, Snavely DB, Hargreaves RJ, Szegedi A, Lines C, Reines SA. Addition of an Nk1 receptor antagonist to an Ssri did not enhance the antidepressant effects of Ssri monotherapy: results from a randomized clinical trial in patients with major depressive disorder. Hum Psychopharmacol. 2014;29:568–77.

122. Chen YM, Huang XM, Thompson R, Zhao YB. Clinical features and efficacy of escitalopram treatment for geriatric depression. J Int Med Res. 2011;39(5):2011.

123. Clayton AH, Croft H, Horrigan JP, Wightman DS, Krishen A, Richard NE, Modell JG. Bupropion extended release compared with escitalopram: Effects on sexual functioning and antidepressant efficacy in 2 randomized, double-blind, placebo-controlled studies. J Clin Psychiatry. 2006;67:736–46.

124. Coleman CC, Cunningham LA, Foster VJ, Batey SR, Donahue RM, Houser TL, Ascher JA. Sexual dysfunction associated with the treatment of depression: a placebo-controlled comparison of bupropion sustained release and sertraline treatment. Ann Clin Psychiatry. 1999;11:205–15.

125. Croft H, Settle E, Houser T, Batey SR, Donahue RM, Ascher JA. A placebo-controlled comparison of the antidepressant efficacy and effects on sexual functioning of sustained-release bupropion and sertraline. Clin Ther. 1999; 21:643–58.

126. Doogan DP, Langdon CJ. A double-blind, placebo-controlled comparison of sertraline and dothiepin in the treatment of major depression in general practice. Int Clin Psychopharmacol. 1994;9:95–100.

127. Fabre LF, Clayton AH, Smith LC, Goldstein I, Derogatis LR. The Effect of Gepirone-ER in the Treatment of Sexual Dysfunction in Depressed Men. J Sex Med. 2012, 9(3):March.

128. Golden RN, Nemeroff CB, McSorley P, Pitts CD, DubÇ EM. Efficacy and tolerability of controlled-release and immediate-release paroxetine in the treatment of depression. J Clin Psychiatry. 2002;63:577–84.

129. Hunter AM, Leuchter AF, Cook IA, Abrams M. Brain functional changes (QEEG cordance) and worsening suicidal ideation and mood symptoms during antidepressant treatment. Acta Psychiatr Scand. 2010;122:461–9.

130. Lepola UM, Loft H, Reines EH. Escitalopram (10–20 mg/day) is effective and well tolerated in a placebo-controlled study in depression in primary care. Int Clin Psychopharmacol. 2003;18:211–7.

131. Leuchter AF, Morgan M, Cook IA, Dunkin J, Abrams M, Witte E. Pretreatment neurophysiological and clinical characteristics of placebo responders in treatment trials for major depression. Psychopharmacology (Berl). 2004;177:15–22.

132. Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, Hróbjartsson A, Mann H, Dickersin K, Berlin JA, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013; 158(3):200–7.

133. Mao JJ, Xie SX, Zee J, Soeller I, Li QS, Rockwell K, Amsterdam JD. Rhodiola rosea versus sertraline for major depressive disorder: A randomized placebo-controlled trial. Phytomedicine. 2015;22:394–9.

134. Mathews M, Gommoll C, Nunez R, Khan A. Efficacy and safety of vilazodone 20 and 40 Mg in major depressive disorder: a randomized, double-blind, placebo-controlled trial. Int Clin Psychopharmacol. 2015;30:67–74.

135. Mendels J, Kiev A, Fabre LF. Double-blind comparison of citalopram and placebo in depressed outpatients with melancholia. Depress Anxiety. 1999; 9:54–60.

136. Montgomery SA, Rasmussen JG, Lyby K, Connor P, Tanghõj P. Dose response relationship of citalopram 20 mg, citalopram 40 mg and placebo in the treatment of moderate and severe depression. Int Clin Psychopharmacol. 1992;6 Suppl 5:65–70.

137. MY-1045/BRL-029060/1 (PAR128) (Unpublished Data) GlaxoSmithKline. A Multicenter, Randomized, Double-Blind, Placebo-Controlled Comparison of Paroxetine and Fluoxetine in the Treatment of Major Depressive Disorder. http://www.gsk-clinicalstudyregister.com/files2/2821.pdf 2005, 1:1–5.

138. MY-1043/BRL-029060/115 (Unpublished Data) GlaxoSmithKline. A multicenter, randomized, double-blind, placebo-controlled comparison of paroxetine and fluoxetine in the treatment of major depressive disorder. http://www.gsk-clinicalstudyregister.com/files2/2060.pdf 2005, 175:1–5.

139. Rath A, Ngwabyt S, Durisic S, Garattini S, Neugebauer EAM, Laville M, Jakobsen JC, Kubiac C, DeMotes-Mainard J, Gluud C: Specific barriers to the conduct of randomised clinical trials within rare diseases - the Europeean Clinical Research Infrastructure (ECRIN) perspective. TRIALS, to be submitted 2015.

140. AstraZeneca. AZD7268 Safety and Tolerability Study. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000- [cited 2015 Jun 19]. https://clinicaltrials.gov/ct2/show/results/NCT01020799?term= NCT01020799&rank=1. In.; 2012.

141. Nierenberg, Greist JH, Mallinckrodt CH, Prakash A, Sambunaris A, Tollefson GD, Wohlreich MM. Duloxetine versus escitalopram and placebo in the treatment

142. PAR 01-001(Unpublished Data). Study report from U.S. Food and Drug Administration. In.; 1992.

143. PAR 09 (Unpublished Data). Study report from U.S. Food and Drug Administration. In.; 1992.

144. Nyth aL, Gottfries CG, Lyby K, Smedegaard-Andersen L, Gylding-Sabroe J, Kristensen M, Refsum HE, Ofsti E, Eriksson S, Syversen S. A controlled multicenter clinical study of citalopram and placebo in elderly depressed patients with and without concomitant dementia. Acta Psychiatr Scand. 1992;86:138–45.

145. PAR 29060.07.001 (Unpublished Data) GlaxoSmithKline. A Double-Blind Comparison of Paroxetine, Amitriptyline, and Placebo in Inpatients with Major Depressive Disorder with Melancholia. http://www.gsk-clinicalstudyregister.com/files2/1866.pdf 2008:1–4.

146. Perahia DGS, Wang F, Mallinckrodt CH, Walker DJ, Detke MJ. Duloxetine in the treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. Eu Psychiatry. 2006;21:367–78.

147. Peselow ED, Lautin A, Wolkin A. The dexamethasone suppression test and response to placebo. J Clin Psychopharmacol. 1986;6(5):1986.

148. Porro V, Fiorenzioni S, Menga C, De CA, Bertolino A. Single-blind comparison of the efficacy of fluvoxamine versus placebo in patients with depressive syndrome. Curr Therapeutic Res. 1988;43(4):1988.

149. SCT-MD-26 (Unpublished Data) Forest Research Institute. Two-week double-blind placebo controlled study of escitalopram in the treatment of severe major depression. 2005, 135:2–4.

150. SCT-MD-27 (Unpublished Data) Forest Research Institute. A double-blind flexible dose comparison of escitalopram sertraline and placebo in the treatment of major depressive disorder. 2005, 91:5–7.

151. SCT-MD-35 (Unpublished Data) Forest Research Institute. Fixed-dose comparison of escitalopram combination in adult patients with major depressive disorder. 2007:1–28.

152. Walczak DD, Apter JT, Halikas JA, Borison RL, Carman JS, Post GL, Patrick R, Cohn JB, Cunningham LA, Rittberg B, et al. The oral dose-effect relationship for fluvoxamine: a fixed-dose comparison against placebo in depressed outpatients. Ann Clin Psychiatry. 1996;8:139–51.

153. Zitman FG, CouvÇe JE. Chronic benzodiazepine use in general practice patients with depression: an evaluation of controlled treatment and taper-off: report on behalf of the Dutch Chronic Benzodiazepine Working Group. Brit J Psychiatry. 2001;178:317–24.

154. Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, Ioannidis JP, Straus S, Thorlund K, Jansen JP, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. Ann Intern Med. 2015;162(11):777–84.

155. 29060/785 (Unpublished Data) GlaxoSmithKline. A double-blind, placebo-controlled, fixed-dosage study comparing the efficacy and tolerability of paroxetine CR and citalopram to placebo in the treatment of Major Depressive Disorder with anxiety. http://www.gsk-clinicalstudyregister.com/files2/2202.pdf 2005:1–5.

156. Cornelius JR, Salloum IM, Ehler JG, Jarrett PJ, Cornelius MD, Perel JM, Thase ME, Black A. Fluoxetine in depressed alcoholics. A double-blind, placebo-controlled trial. Arch Gen Psychiatry. 1997;54:700–5.

157. Goldstein DJ, Lu Y, Detke MJ, Wiltse C, Mallinckrodt C, Demitrack M. Duloxetine in the treatment of depression. J Clin Psychopharmacol. 2004;24:389–99.

158. Goldstein DJ, Ph D, Mallinckrodt C, Ph D, Lu Y, Ph D, Demitrack MA. Duloxetine in the treatment of major depressive disorder: a double-blind clinical trial. J Clin Psychiatry. 2002;63(3):225–31.

159. Kaddurah-Daouk R, Boyle SH, Matson W, Sharma S, Matson S, Zhu H, Bogdanov MB, Churchill E, Krishnan RR, Rush aJ, et al. Pretreatment metabotype as a predictor of response to sertraline or placebo in depressed outpatients: a proof of concept. Translat Psychiatry. 2011;1:e26.

160. Kennedy S. Flibanserin: initial evidence of efficacy on sexual dysfunction, in patients with major depressive disorder. J Sex Med. 2010;7(10):3449–59.

161. Dubini A, Bosc M, Polin V. Noradrenaline-selective versus serotonin-selective antidepressant therapy: differential effects on social functioning. J Psychopharmacol. 1997;11:S17–23.

162. Tollefson GD, Bosomworth JC, Heiligenstein JH, Potvin JH. A double-blind, placebo-controlled clinical trial of fluoxetine in geriatric patients with major depression. Int Psychogeriatr. 1995;1:89–104.

Jakobsen *et al. BMC Psychiatry* (2017) 17:58

Page 28 of 28

163. Dunlop SR, Dornseif BE, Wernicke JF, Potvin JH. Pattern analysis shows beneficial effect of fluoxetine treatment in mild depression. Psychopharmacol Bull. 1990;26(2):1.

164. Schmitz JM, Averill P, Stotts AL, Moeller FG, Rhoades HM, Grabowski J. Fluoxetine treatment of cocaine-dependent patients with major depressive disorder. Drug Alcohol Depend. 2001;63:207–14.

165. 89306 (Unpublished Data). Study report from U.S. Food and Drug Administration. In.; 1998.

166. Wakelin JS. Fluvoxamine in the treatment of the older depressed patient; double-blind, placebo-controlled data. Int Clin Psychopharmacol. 1986;1: 221–30.

167. 29060/448 (Unpublished Data) GlaxoSmithKline. A double-blind, placebo controlled trial to evaluate the clinical effects of immediate release paroxetine and modified release paroxetine in the treatment of major depression. URL: https://www.gsk-clinicalstudyregister.com/files2/2021.pdf. 2005.

168. NKF100096 (Unpublished Data) GlaxoSmithKline. A randomised, double-blind, double-dummy, parallel-group, placebo-controlled, forced dose titration study evaluating the efficacy and safety of a new chemical entity (NCE) and paroxetine in subjects with major depressive disorder (MDD). 2007:4–7.

169. 89303 (Unpublished Data) Forest Laboratories Inc. Study report from U.S. Food and Drug Administration. In.; 1998.

170. Cohn JB, Wilcox CS. Paroxetine in major depression: a double-blind trial with imipramine and placebo. J Clin Psychiatry. 1992;53(Suppl):52–6.

171. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and cox regression. Am J Epidemiol. 2007;165(6):710–8.

172. Tollefson GD, Holman SL. Analysis of the Hamilton Depression Rating Scale factors from a double-blind, placebo-controlled trial of fluoxetine in geriatric major depression. Int Clin Psychopharmacol. 1993;8:253–9.

173. CIT-MD-03 (Unpublished Data) Forest Laboratories Inc. A randomized double-blind placebo-controlled trial of citalopram in depressed patients at least 75 years of age. 2005:1–4.

174. Bech AT. An inventory for measuring depression. Arch Gen Psychiatry. 1961; 4:561–71.

175. SND103288 (Unpublished Data) GlaxoSmithKline. A ten-week, multicentre, randomised, double-blind, placebo-and active-controlled, parallel-group, flexible-dose study evaluating the efficacy, safety and tolerability of GSK372475, a new chemical entity (NCE), or paroxetine compared to placebo in adult subjects diagnosed with major depressive disorder. URL: http://www.gsk-clinicalstudyregister.com/files2/21101.pdf. 2008:1–9.

176. WELL AK130926 (Unpublished Data) GlaxoSmithKline. A multicenter, double-blind, randomized, placebo-controlled comparison of the effects on sexual functioning of extended-release bupropion hydrochloride (300–450 mg) and escitalopram (10–20 mg) in outpatients with moderate to severe major depression over an eight-week treatment period. URL: http://www.gsk-clinicalstudyregister.com/files2/20466.pdf. 2007.

177. Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. Stat Med. 2004;23(9):1351–75.

178. Forest Laboratories Inc: SCT-MD-01. Unpublished 2001, Study report from FDA.

179. 29060/810 (Unpublished Data) GlaxoSmithKline. A randomized, double-blind, parallel-group, placebo-controlled study evaluating efficacy and safety of paroxetine controlled release (12.5 and 25 mg/day) versus placebo in patients with major depressive disorder. URL: https://www.gsk-clinicalstudyregister.com/files2/2204.pdf. 2005.

180. Forest Laboratories. Safety and efficacy of vilazodone in major depressive disorder (VLZ-MD-01). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000- [cited 2015 Jun 19]. Available from: http://clinicaltrials.gov/show/NCT01473381 URL: https://clinicaltrials.gov/ct2/show/results/NCT01473381?term=NCT01473381&rank=1. In.; 2014.

181. SCT-MD-02 (Unpublished Data) Forest Laboratories Inc. Study report from U.S. Food and Drug Administration. In.; 2002.

182. Broich K. Committee for Medicinal Products for Human Use (CHMP) assessment on efficacy of antidepressants. Eur Neuropsychopharmacol. 2009;19(5):305–8.

183. Hunt SM, McKenna SP. The QLDS: a scale for the measurement of quality of life in depression. Health Policy. 1992;22(3):307–19.

184. Jakobsen JC, Gluud C. The necessity of randomized clinical trials. Br J Med Res. 2013;3(4):1453–68.

185. Rimer J, Dwan K, Lawlor Debbie A, Greig Carolyn A, McMurdo M, Morley W, Mead Gillian E. Exercise for depression. Coch Database Syst Rev. 2012;7: CD004366. pub004366.

186. Leucht S, Fennema H, Engel R, Kaspers-Janssen M, Lepping P, Szegedi A. What does the HAMD mean? J Affect Disord. 2013;148(2–3):243–8.

187. Moncrieff J, Kirsch I. Empirically derived criteria cast doubt on the clinical significance of antidepressant-placebo differences. Contemp Clin Trials. 2015;43:60–2.

188. Furukawa TA, Cipriani A, Atkinson LZ, Leucht S, Ogawa Y, Takeshima N, Hayasaka Y, Chaimani A, Salanti G. Placebo response rates in antidepressant trials: a systematic review of published and unpublished double-blind randomised controlled studies. Lancet Psychiatry. 2016;3(11):1059–66.

189. Hrobjartsson A, Gotzsche PC. Placebo interventions for all clinical conditions. Cochrane Database Syst Rev. 2010;1:CD003974.

190. Hrobjartsson A, Emanuelsson F, Skou Thomsen AS, Hilden J, Brorson S. Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. Int J Epidemiol. 2014;43(4):1272–83.

191. Hróbjartsson A, Thomsen ASS, Emanuelsson F, Tendal B, Hilden J, Boutron I, Ravaud P, Brorson S. Observer bias in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. BMJ. 2012;344:e1119.

192. Hróbjartsson A, Thomsen ASS, Emanuelsson F, Tendal B, Hilden J, Boutron I, Ravaud P, Brorson S. Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. CMAJ. 2013;185(4):E201–11.

193. Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. Coch Database Syst Rev. 2012;12:MR000033.

194. Savovic J, Jones HE, Altman DG, Harris RJ, Juni P, Pildal J, Als-Nielsen B, Balk EM, Gluud C, Gluud LL, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. Ann Intern Med. 2012;157(6):429–38.

195. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA. 1995;273(5):408–12.

196. Sutton AJ, Duval SJ, Tweedie RL, Abrams KR, Jones DR. Empirical assessment of effect of publication bias on meta-analyses. BMJ. 2000;320(7249):1574–7.

197. Garattini S, Jakobsen JC, Wetterslev J, Bertele' V, Banzi R, Rath A, Neugebauer EAM, Laville M, Masson Y, Hivert V et al. Evidence-based clinical practice: Overview of threats to the validity of evidence and how to minimise them. Eur J Internal Med. 2016;32:13-21.

198. Krogh J, Hjorthoj CR, Jakobsen JC, Lindschou J, Kessing LV, Nordentoft M, Gluud C. DEPERROR: Risks of systematic errors in drug and non-drug randomized clinical trials assessing intervention effects in patients with unipolar depression. J Affect Disord. 2015;179:121–7.

199. Ragland DR. Dichotomizing continuous outcome variables: dependence of the magnitude of association and statistical power on the cutpoint. Epidemiol. 1992;3(5):434–40.

200. Dawson NV, Weiss R: Dichotomizing continuous variables in statistical analysis. A practice to avoid. Med Decis Making 2012, 32(225).

201. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ. 2006;332(7549):1080.

202. Kirsch I, Moncrieff J. Clinical trials and the response rate illusion. Contemp Clin Trials. 2007;28(4):348–51.

203. Leucht S, Hierl S, Kissling W, Dold M, Davis JM. Putting the efficacy of psychiatric and general medicine medication into perspective: review of meta-analyses. Br J Psychiatry. 2012;200(2):97–106.

204. Brok J, Thorlund K, Wetterslev J, Gluud C. Apparently conclusive meta-analyses may be inconclusive–Trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses. Int J Epidemiol. 2009;38(1):287–98.

205. Savović J, Jones H, Altman D, Harris R, Juni P, Pildal J, Nielsen B, Balk E, Gluud C, Gluud L, et al. Influence of reported study design characteristics on intervention effect estimates from randomized controlled trials: combined analysis of meta-epidemiologic studies. Health Technol Assess. 2012;16(35):1–82.

206. Andrews P, Thomson JA, Amstadter A, Neale M. Primum Non Nocere: An Evolutionary Analysis of Whether Antidepressants Do More Harm than Good. Front Psychol. 2012, 3(117).

207. Gøtzsche PC: Deadly Medicines and Organised Crime: How Big Pharma Has Corrupted Healthcare: Radcliffe Medical Press ltd; 2013.

208. Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, Watanabe N, Nakagawa A, Omori IM, McGuire H, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. Lancet. 2009;373(9665):746–58.