

# Is the efficacy of selective serotonin reuptake inhibitors influenced by adverse events?

## Meta-regression and mediation analysis of placebo-controlled trials

Michael Barth, Levente Kriston, Swaantje Klostermann, Corrado Barbui,  
Andrea Cipriani and Klaus Linde

**Keywords:** major depression, drug therapy, selective serotonin reuptake inhibitor,  
adverse events, meta-analysis, meta-regression analysis, mediation analysis

**Word count (abstract):** 210 words  
**Word count (main text):** 3402 words  
**Figures:** 3  
**Tables:** 2  
**References:** 32

## **Summary**

### **Background**

It has been suggested that the efficacy of antidepressants has been overestimated in clinical trials due to unblinding of drug treatments by adverse events.

### **Aims**

To investigate the association between adverse events and the efficacy of selective serotonin reuptake inhibitors (SSRIs).

### **Method**

The literature was searched to identify randomized, double-blind, placebo-controlled trials of SSRIs in the treatment of major depression. Efficacy outcomes were response to treatment and change of depressive symptoms. Reporting of adverse events was used as indicator of tolerability. Random effects meta-analyses were used to calculate pooled estimates. Meta-regression analyses were performed to investigate the association between adverse events and efficacy. Potential mediation was investigated with the Baron and Kenny approach.

### **Results**

68 trials with 17,646 subjects were included in the analyses. In meta-analysis SSRIs were superior to placebo regarding efficacy (OR 1.62; 95%CI 1.51, 1.72). More subjects allocated to SSRIs reported adverse events than subjects receiving placebo (OR 1.73; 95%CI 1.58, 1.89). Meta-regression analyses did not find an association between adverse events and efficacy ( $p=0.439$ ). There was no indication of adverse events mediating the effect of SSRI treatment.

### **Conclusions**

Our results do not support, but also do not unequivocally disprove, the hypothesis that adverse events lead to an overestimation of the effect of SSRIs over placebo.

### **Declarations of interest**

None

## Introduction

Selective serotonin reuptake inhibitors (SSRIs) are the common standard of pharmacological treatment of depressive disorders in adults and are recommended in major clinical practice guidelines.<sup>1-3</sup> However, some critics doubt that SSRIs and other antidepressants have clinically relevant benefits over placebo,<sup>4-6</sup> because drug-placebo differences become smaller if unpublished trials are included in meta-analyses.<sup>7</sup> Some researchers argued that even the remaining small differences between drugs and placebo may be attributable to the unblinding of active treatment, because patients may find out whether they receive drug or placebo due to the side effects of the active drug.<sup>5</sup>

Unblinding could increase placebo effects in the true treatment group through increased expectancy and leading to an overestimation of the efficacy of the active drug.<sup>5</sup> When this is true one would expect that trials showing larger differences regarding adverse events between drug and placebo also show larger differences regarding improvement. An older meta-analysis indeed found a strong association between efficacy and adverse effects but the results were based on only six trials.<sup>8</sup> To our knowledge no current systematic review addressing this issue exists. Therefore, we aimed to investigate whether there is evidence of an association between adverse effects and efficacy in a larger set of trials comparing SSRIs with placebo.

## Methods

### Literature search and study selection

Potentially relevant articles were searched in the Cochrane Register of Controlled Trials (CENTRAL) and PubMed (last update on 2 August 2013). In addition, we screened reference lists of relevant reviews and included trials. We then searched clinical trial registries of drug companies (see supplementary information). Two independent reviewers (MB, SK) screened titles and abstracts of retrieved citations and excluded clearly irrelevant reports. Subsequently, the full texts of all potentially relevant papers were obtained and checked against the inclusion criteria by two reviewers. Disagreements were resolved by discussion with a third reviewer (KL).

To be included in our review studies, had to be double-blind, placebo-controlled, randomized trials with at least 10 patients per study group and an active treatment period (post-randomization) between 4 and 13 weeks. Participants were adults suffering from an acute (single or recurrent) episode of moderate to severe major depression (at least 14 points on the Hamilton Rating Scale for Depression (HRSD)<sup>9</sup> or 20 points on the Montgomery–Åsberg Depression Rating Scale (MADRS)<sup>10</sup>) according to the criteria of Diagnostic and Statistical Manual of Mental Disorders (DSM), version III or later. Studies with a majority of patients over 65 years (median>65) and under 18 years of age were excluded. Eligible treatments were SSRIs (citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, sertraline, vilazodone) or placebo, prescribed in adequate dosage according to the American Psychiatric Association Guideline<sup>1</sup> as monotherapy. The number or proportion of patients experiencing a symptom scale score reduction of at least 50% on the HRSD or the MADRS had to be reported or imputable from available score data using the normal distribution assumption. Furthermore, the number or proportion of patients experiencing adverse events or adverse effects had to be reported.

## **Data extraction, major outcomes and assessment of the risk of bias**

Two independent reviewers (MB, SK) extracted data on subjects, methods, intervention details and outcomes from the included studies using a standardized form. Data from publications were supplemented by clinical study reports, clinical trial registers and publicly available reports of the US Food and Drug Administration (FDA), the German Institute for Quality and Efficacy in Healthcare (IQWiG), and the UK National Institute for Health and Care Excellence (NICE).

The primary efficacy outcome of our review was response after treatment. Response was defined as the number of patients who had a reduction of at least 50% in symptoms' severity between baseline and endpoint on a standardized rating scale. When trials reported efficacy results according to more than one rating scale, the rating scale chosen as primary outcome by the study authors was used in the analysis; otherwise the HRSD results were prioritized.

Subjects who were randomized but not included in the intention-to-treat-analysis were counted as non-responders. As a secondary efficacy outcome, we extracted mean change of depression scores from baseline (mean score at end of treatment minus mean score baseline) with its respective standard deviation. When only standard errors or p-values were reported, standard deviations were calculated according to the Cochrane Handbook.<sup>11</sup> The primary safety outcome was the number or proportion of subjects experiencing at least one adverse event. If data on adverse events were not reported, reports on adverse effects were used. In case of discrepancies between journal publications and study reports (or summaries) from pharmaceutical manufacturer registers, data from the latter was used.

The risk of bias in included studies was assessed with the Cochrane Collaboration's tool<sup>12</sup> for the following domains: a) adequacy of the random sequence generation, b) concealment of allocation, c) blinding of the participants, personnel, and outcome assessors, and d) incomplete outcome data (application of the intention-to-treat principle). We did not assess the risk of bias due to selective reporting and other sources of bias as our inclusion criteria warranted that all studies reported the outcome data needed for analysis. The risk of bias was categorized into low, unclear or high risk for each domain. Overall risk of bias was considered high if one or more domains were rated as high risk and low if at least three domains were rated as low risk and none was rated high.

## Data synthesis

To summarize the findings of individual studies we calculated odds ratios (OR) and their corresponding 95% confidence intervals (CI) for response and for the number of patients reporting at least one adverse event. Standardized mean differences (SMD) were calculated for change from baseline. An OR above one indicates more responders and more patients with adverse events in the treatment group, respectively. A negative SMD indicates more symptom reduction in the treatment group. To ensure that each patient appeared only once in the meta-analysis, we pooled intervention groups and split placebo groups according to the Cochrane Handbook.<sup>11</sup>

We utilized correlation analyses (Pearson correlation coefficients, inverse variance weighted analysis) to investigate the relationship between adverse events and efficacy within and between the groups. We used the logarithm of odds of response and adverse events, respectively, to measure the change in symptom severity from baseline effect size (change from baseline mean divided by its standard deviation) for within group analyses. Logarithm of OR of response and adverse events (and SMD of change from baseline) were used for the analyses between groups. Taking the log of odds and odds ratios enabled the fitting of linear models (correlation, regression) with the assumption of normally distributed residuals.

We performed random effects meta-analyses with inverse variance weighting<sup>13</sup> for the outcomes response, number of patients with adverse events and change from baseline. Between-study heterogeneity was assessed using  $\tau^2$ ,  $I^2$  and the  $\chi^2$  test.

We used inverse variance weighted random effects meta-regression<sup>14</sup> to investigate whether there was a relationship between the logarithmic OR (log OR) for the number of patients with adverse events and the log OR for response or the SMD from change from baseline, respectively. We also investigated whether potential confounders had an influence on this

relationship, including them and their interaction with the log OR of patients with adverse events in a stepwise manner. Pre-defined variables were mean severity of depression (low vs high defined as HRSD below or above 25 or MADRS score below or above 35)<sup>3,15</sup>; dosing of medication (classified according to Gartlehner et al<sup>16</sup>; overall risk of bias; publication status; response and adverse event rates (log odds), respectively, in the placebo group; and trial size (median split). As post hoc analyses, we added reporting of response data (vs. needing to impute response from continuous data), adverse events specification (treatment-related vs. other) and the probability to be assigned to the placebo group.<sup>17</sup> To minimize multicollinearity, interaction terms were built from centered variables.

To investigate whether adverse events mediate the relationship between treatment and efficacy, we used a mediation framework.<sup>18</sup> These analyses were performed on the study arm level. According to this approach mediation exists when four conditions are met: firstly (pathway c) - the predictor (in this case treatment) must be significantly related to the outcome variable (response log odds); secondly (pathway a) - the predictor must also be significantly related to the potential mediator (adverse events log odds); thirdly (pathway b) - the mediator must be significantly related to the outcome when the effect of the predictor on the outcome is controlled for; and fourthly (pathway c') - the relationship between predictor and outcome must be decreased (lower than in pathway c) when controlling for the mediator.

If the predictor remains significant when the mediator is controlled for, the mediation is considered partial. When controlling for the mediator renders the predictor non-significant, mediation is considered complete. Statistical significance was set to  $P < 0.05$ . To avoid that some true mediation effects are missed<sup>19</sup> indirect effects were calculated<sup>20,21</sup> and tested for significance.<sup>22,23</sup> Mediation analyses were conducted with random effects model (via restricted information maximum likelihood) and inverse variance weighting.

Meta-analyses were performed using Review Manager 5.2 (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen). Meta-regression and mediation analyses were

conducted with IBM SPSS Statistics for Windows, Version 21.0 (IBM Corporation, Armonk, NY) using a macro by Wilson.<sup>24</sup>

## **Results**

### **Literature search and selection process**

The literature search identified 1,002 potentially relevant references (see figure 1). 856 records were excluded after the screening of titles and abstracts and a further 146 after checking of full texts. A total of 68 studies published in 95 reports or publications were included in the systematic review (see supplementary information). The most frequent reason for exclusion of studies was lack of reporting of proportion of patients with at least one adverse event.

### **Characteristics of included studies**

The 68 studies included a total of 17,646 subjects (10,376 allocated to active treatment and 7,270 to placebo); the median number of subjects per study was 252 (range: 25 to 877). An overview of study characteristics is given in table 1. Characteristics of the individual studies are summarized in the appendix (see supplementary table 1). Thirteen trials (19%) were completely unpublished and five (7%) were only published as congress abstracts. Data from ten additional trials (15%) was not published in stand-alone journal articles but included in publications pooling primary data from at least two trials. Fifty-four studies were multi-arm trials which compared different active agents or different dosages of the same agent with placebo. The overall risk of bias was considered high for 33 studies, unclear for 27 and low in eight studies.

### **Availability of outcome data and basic findings**

The number of responders was reported in 48 studies and was estimated from continuous data in 20 studies. Means for symptom severity change from baseline were reported in 66 studies, but for 48 studies standard deviations had to be imputed. The number of patients with “adverse



events” or “any adverse event” was reported by 43 studies, while the remaining 25 studies reported only the number of patients with adverse effects (i.e., adverse events related to the treatment, such as “treatment emergent adverse events/symptoms” or “side effects”). The proportion of responders ranged from 14% to 74% (median 48%) in groups receiving SSRIs and from 8% to 59% (median 38%) in placebo groups, the standardized mean severity change from baseline within groups ranged from 0.80 to 2.83 (median 1.40) and 0.38 to 2.05 (median 1.10), respectively. The proportion of patients reporting adverse events ranged from 16% to 98% (median 80%) under SSRI treatment and from 17% to 94% (median 70%) under placebo.

There was no statistically significant correlation in the treatment group between response and adverse events (inverse variance weighted analysis using log odds transformed values, Pearson correlation coefficients). In contrast, there was a moderate correlation between the effect size’s mean severity change from baseline and adverse events in the treatment group ( $r=0.280$ ,  $p=0.009$ ). No statistically significant correlations between response and adverse events were found in placebo groups.

## **Meta-analysis**

In the random effect meta-analysis SSRIs showed significantly higher response rates than placebo (OR 1.62; 95%CI 1.51, 1.72; see table 2). There was no evidence of between-study heterogeneity ( $p=0.49$  and  $I^2=0\%$ ). There was no indication that effects over placebo varied between different agents ( $p=0.98$  and  $I^2=0\%$  in the test for subgroup differences). SSRIs were also found to be superior to placebo when symptom severity change from baseline was used as outcome measure (SMD -0.27; 95%CI -0.31, -0.23). There was some between-study heterogeneity ( $p<0.001$ ;  $I^2=33\%$ ) but no evidence of subgroup differences according to agents ( $p=0.33$ ;  $I^2=13\%$ ). Patients allocated to SSRIs reported significantly more adverse events than patients receiving placebo (OR 1.73; 95%CI 1.58, 1.89; table 2). There was low between-study heterogeneity ( $p=0.02$ ;  $I^2=28\%$ ) and evidence for differences between single agents ( $p<0.01$ ;  $I^2=68\%$ ).

## Meta-regression analysis

We did not find any statistically significant correlation between response over placebo (log OR response) and group differences regarding adverse events. Similarly, in the meta-regression there was no significant association between efficacy (log OR response) and tolerability (log OR adverse events) of SSRIs in comparison to placebo ( $\beta=0.064$ ;  $SE=0.083$ ;  $p=0.439$ ; see figure 2). We found no statistically significant correlation between SMD for change from baseline values and adverse events. In the meta-regression analysis with adverse events as predictor and mean severity change from baseline as dependent variable no statistically significant effect was found ( $\beta=0.062$ ;  $SE=0.045$ ;  $p=0.167$ ). Adding confounders to the model did not provide evidence that the association of adverse and beneficial effects was suppressed or varied across subgroups except for a statistically significant moderating effect of trial size ( $p=0.04$ ) suggesting that the investigated association may be more pronounced in small rather than in large trials (see supplementary table 2).

## Mediation analysis

SSRI treatment was significantly associated with better treatment response in pathway c ( $\beta=0.499$ ,  $SE=0.071$ ,  $p<0.001$ ; figure 3) and with higher rates of subjects with adverse events in pathway a ( $\beta=0.663$ ,  $SE=0.136$ ,  $p<0.001$ ). In pathway b, no significant association was found between adverse events and outcome ( $\beta= -0.036$ ,  $SE=0.041$ ,  $p=0.384$ ). The direct relationship between SSRI treatment and treatment response under control of adverse events (pathway c') was statistically significant ( $\beta=0.524$ ,  $SE=0.077$ ,  $p<0.001$ ). By contrast, the indirect effect of mediation was not statistically significant ( $\beta= -0.024$ ,  $SE=0.028$ ,  $p=0.402$ ). As one of the conditions for mediation (association of mediator with the outcome under control of the predictor) remained unfulfilled, no evidence of the mediating role of adverse events between treatment and response was found. These results were confirmed when change from baseline was used as the outcome variable, when potential confounders were included in the analyses

and when analyses were restricted to the SSRI tested most often (see supplementary tables 4 and 5).

## Discussion

In the present systematic review of 68 studies we investigated the potential association between efficacy and adverse events during SSRI-treatment. In the preliminary meta-analysis we found small beneficial effects of SSRI compared to placebo in terms of efficacy, with a higher proportion of subjects reporting adverse events. The meta-regression analyses, however, did not show any significant association between adverse events and efficacy, even after controlling for clinical (i.e. severity of depression, medication dosage) and methodological variables (i.e. overall risk of bias, publication status, trial size). Consistently, the mediation analysis confirmed these results suggesting no mediation effect between treatment itself (SSRI or placebo) and efficacy due to adverse events.

In terms of efficacy, our findings are similar to previous meta-analyses.<sup>7,25,26</sup> Many placebo-controlled trials on SSRI are unpublished and if these are included in the analyses, effects over placebo are small. However, for all SSRIs the effect over placebo was statistically significant and consistent with effect estimates showing no or little statistical heterogeneity.

We did not find a statistically significant association between adverse effects and efficacy. This is in contrast to the findings of Greenberg.<sup>8</sup> In their meta-analysis of fluoxetine compared to placebo the authors reported a strong correlation between adverse events and efficacy ( $r = 0.85$  for clinician-rated and  $r=0.96$  for patient-rated efficacy measures). On the basis of these findings, Greenberg et al. supposed that outcome scores for patients allocated to active drug may be amplified when study participants become aware of treatment by experiencing adverse events. However, the correlation analyses by Greenberg et al. were based on only four to six clinical trials. We are not aware of any published meta-regression and mediation analyses on this topic. The main strength of meta-regression is the possibility in analyzing moderators and weighting for precision of estimates (with large precise studies obtaining a larger weight).

However, to investigate the indirect effects from treatment to efficacy mediated by adverse events, arm-based analyses are needed (no between-group outcomes (e.g. OR), but rather within-group ones (e.g. odds)). Mediation solves this problem by comparing the treatment-arms of different studies to each other. We considered this kind of analysis to be the best possible way to investigate our question about the mediating effect of adverse events in SSRI-treatment within a meta-analytical framework. Again neither meta-regression nor mediation analysis provided any hints that adverse effects have any influence on efficacy.

When interpreting our findings three important limitations have to be considered. First, as previous analyses<sup>8</sup> we had to use adverse events or adverse effects as a surrogate for unblinding and resulting changes in expectations. Only a tiny proportion of drug trials assess blinding. This applies to our data set, to trials in affective disorders or schizophrenia in general<sup>27</sup> and a random sample of trials from the medical literature.<sup>28</sup> Therefore, it was impossible to investigate the association between unblinding and efficacy directly. Furthermore, assessing unblinding is problematic and the recommendation to report related results has been deleted in the second revision of the CONSORT guidelines for reporting randomized controlled trials.<sup>29</sup> Similarly, to the best of our knowledge validated methods to assess changes in patient expectations during a trial do not exist.

Second, we had to use the number of patients reporting at least one adverse event or adverse effect to investigate the association between adverse events and efficacy. Specific adverse effects (e.g. sexual dysfunction) are likely to have a higher impact on unblinding and related expectations than non-specific adverse events and adverse effects. Minor adverse effects might lead to unblinding (leading to an overestimation of treatment effects), more severe adverse effects may lead to drop-out (possibly decreasing group differences depending on the method for replacing missing values). However, methods to assess specific adverse effects and their reporting are extremely variable.<sup>30</sup> In our study population the proportion of patients reporting adverse events varied between 16% and 98%. When a high proportion of patients in the active treatment group reported adverse events, the placebo group reported this as well.<sup>31</sup> This strong

association is most likely due to methodological reasons. If any adverse *event* is reported, this leads to higher numbers than reporting of treatment-related adverse *effects* only. Structured interviews with specific symptom lists produce other rates of reported side effects than open questions.<sup>30,32</sup> Furthermore, previous depression and intake of antidepressant drugs in the past can lead to sensitization and symptom provocation.<sup>33</sup> Some studies report specific adverse details in great detail (with highly variable categorizations of groups of side effects), while others provide only overall numbers. This made it impossible to investigate the association between specific adverse effects and efficacy in our data set. The number of patients with at least one adverse event or adverse effect allows straightforward extraction of a defined numerator and denominator (the number of patients randomized) even when trials used variable assessment methods and ensures that a subject is included only once in the analysis.

A third important limitation of our analysis is the use of aggregated data (proportion of patients with adverse events and response, and mean change from baseline). It would be clearly more sensitive and more valid to investigate the potential association and mediation of adverse events and efficacy in individual patient data meta-analysis.<sup>34,35</sup> However, also the use of individual patient data does not resolve the problem that trials use variable methods to assess adverse effects and reliable methods for measuring unblinding are not available.

In conclusion, our results do not support the hypothesis that efficacy of antidepressants over placebo in randomized trials is biased by the unblinding due to adverse effects. These findings were consistent across different efficacy outcomes and statistical methods. However, due to fundamental methodological challenges inherent to the topic investigated, our results should be interpreted with caution and considered only as preliminary evidence. As at least some of these issues could be addressed in ad hoc large trials or individual patient data meta-analyses, such studies would be desirable.

**Michael Barth**, MSc, **Klaus Linde**, MD, Institute of General Practice, Klinikum rechts der Isar, Technische Universität München, Munich, Germany;

**Levente Kriston**, PhD, Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany;

**Swaantje Klostermann**, MPH, Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, University Hospital of Munich (Ludwig-Maximilians-University) Munich, Germany;

**Corrado Barbui**, MD, **Andrea Cipriani\***, MD, WHO Collaborating Centre for Research and Training in Mental Health and Service Evaluation, Department of Public Health and Community Medicine, Section of Psychiatry, University of Verona, Verona, Italy; \*Department of Psychiatry, University of Oxford, Warneford Hospital, UK

**Correspondence:**

Michael Barth, Institute for General Practice, Klinikum Rechts der Isar, Technische Universität München, Orleansstr. 47, 81667 München, Germany.

Email: michael.barth@tum.de

## References

1. Work Group on Major Depressive Disorder. *Practice guideline for the treatment of patients with major depressive disorder*. American Psychiatric Association (APA); 2010: (<http://psychiatryonline.org/pdfaccess.ashx?ResourceID=243261&PDFSource=6>).
2. National Collaborating Centre for Mental Health (NCCMH). *Depression: the Treatment and Management of Depression in Adults (Update)*. NCCMH; 2010.
3. Deutsche Gesellschaft für Psychiatrie, Psychotherapie und Nervenheilkunde (DGPPN). *S3-Leitlinie/Nationale VersorgungsLeitlinie Unipolare Depression Langfassung (1.3th edn.)*. Ärztliches Zentrum für Qualität in der Medizin (ÄZQ); 2012. p. 264.
4. Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; **141**: 781-8.
5. Kirsch I. *The Emperor's New Drugs: Exploding the Antidepressant Myth*. Basic Books, 2009
6. Moncrieff J. Are antidepressants overrated? A review of methodological problems in antidepressant trials. *J Nerv Ment Dis* 2001; **189**: 288-95.
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**: 252-60.
8. Greenberg RP, Bornstein RF, Zborowski MJ, Fisher S, Greenberg MD. A meta-analysis of fluoxetine outcome in the treatment of depression. *J Nerv Ment Dis* 1994; **182**: 547-51.
9. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960; **23**: 56-62.
10. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979; **134**: 382-9.
11. Higgins JPT, Deeks JJ. Selecting studies and collecting data. In *Cochrane Handbook for Systematic Reviews of Interventions* (eds JPT Higgins, S Green): 151-85. John Wiley & Sons, 2008.
12. Higgins JPT, Deeks JJ. Assessing risk of bias in included studies. In *Cochrane Handbook for Systematic Reviews of Interventions* (eds JPT Higgins, S Green): 187-235. John Wiley & Sons, 2008.
13. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**:177-88.
14. Lipsey MW, Wilson DB. *Practical Meta-Analysis (1 ed.)*. SAGE Publications; 2000.
15. Cusin C, Yang H, Yeung A, Fava M. Rating Scales for Depression. In: *Handbook of Clinical Rating Scales and Assessment in Psychiatry and Mental Health* (eds L Baer, MA Blais): Humana Press; 2010.
16. Gartlehner G, Hansen RA, Thieda P, DeVeaugh-Geiss AM, Gaynes BN, Krebs EE, et al. *Comparative Effectiveness of Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression. Comparative Effectiveness Review No. 7*. Agency for Healthcare Research and Quality; 2007
17. Papakostas GI, Fava M. Does the probability of receiving placebo influence clinical trial outcome? A meta-regression of double-blind, randomized clinical trials in MDD. *Eur Neuropsychopharmacol* 2009; **19**:34-40.
18. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986; **51**: 1173-82.

19. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annu Rev Psychol* 2007; **58**: 593-614.
20. Sobel ME. Asymptotic confidence intervals for indirect effects in structural equations models. In *Sociological methodology* (ed S Leinhardt): 290-312. Jossey-Bass, 1982.
21. Sobel ME. Some new results on indirect effects and their standard errors in covariance structure models. In *Sociological Methodology 1986* (ed N Tuma): 159-86. American Sociological Association, 1986.
22. Aroian LA. The probability function of the product of two normally distributed variables. *Ann Math Stat* 1944; **18**: 265-71.
23. MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. *Psychol Methods* 2002; **7**: 83-104.
24. Wilson DB. SPSS, Stata, and SAS macros for performing meta-analytic analyses [cited 2013 05.11.]. (<http://mason.gmu.edu/~dwilsonb/ma.html>).
25. Barbui C, Furukawa TA, Cipriani A. Effectiveness of paroxetine in the treatment of acute major depression in adults: a systematic re-examination of published and unpublished data from randomized trials. *CMAJ* 2008; **178**: 296-305.
26. 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**: e45.
27. Baethge C, Asall OP, Baldessarini RJ. Systematic review of blinding assessment in randomized controlled trials in schizophrenia and affective disorders 2000–2010. *Psychother Psychosom* 2013; **82**:152-160.
28. Bello S, Moustgaard H, Hróbjartsson A. The risk of unblinding was infrequently and incompletely reported in 300 randomized clinical trial publications. *J Clin Epidemiol* 2014; **67**:1059-1069.
29. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; **340**: c869.
30. Rief W, Nestoriuc Y, von Lilienfeld-Toal A, Dogan I, Schreiber F, Hofmann SG, et al. Differences in adverse effect reporting in placebo groups in SSRI and tricyclic antidepressant trials: a systematic review and meta-analysis. *Drug Saf* 2009; **32**: 1041-56.
31. Shedden Mora M, Nestoriuc Y, Rief W. Lessons learned from placebo groups in antidepressant trials. *Philos Trans R Soc Lond B Biol Sci* 2011; **366**: 1879–88.
32. Rief W, Avorn J, Barsky AJ. Medication-attributed adverse effects in placebo groups: implications for assessment of adverse effects. *Arch Intern Med* 2006; **166**: 155-60.
33. Colloca L, Benedetti F. How prior experience shapes placebo analgesia. *Pain* 2006; **124**: 126-33.
34. Fournier JC, DeRubeis RJ, Hollon SD, Dimidjian S, Amsterdam JD, Shelton RC, et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA* 2010; **303**: 47-53.
35. 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**: 572-9.