

# Exploring and accounting for publication bias in mental health: a brief overview of methods

doi:10.1136/eb-2013-101700

## ABSTRACT

**Objective** Publication bias undermines the integrity of published research. The aim of this paper is to present a synopsis of methods for exploring and accounting for publication bias.

**Methods** We discussed the main features of the following methods to assess publication bias: funnel plot analysis; trim-and-fill methods; regression techniques and selection models. We applied these methods to a well-known example of antidepressants trials that compared trials submitted to the Food and Drug Administration (FDA) for regulatory approval.

**Results** The funnel plot-related methods (visual inspection, trim-and-fill, regression models) revealed an association between effect size and SE. Contours of statistical significance showed that asymmetry in the funnel plot is probably due to publication bias. Selection model found a significant correlation between effect size and propensity for publication.

**Conclusions** Researchers should always consider the possible impact of publication bias. Funnel plot-related methods should be seen as a means of examining for small-study effects and not be directly equated with publication bias. Possible causes for funnel plot asymmetry should be explored. Contours of statistical significance may help disentangle whether asymmetry in a funnel plot is caused by publication bias or not. Selection models, although underused, could be useful resource when publication bias and heterogeneity are suspected because they address directly the problem of publication bias and not that of small-study effects.

## INTRODUCTION

Synthesis of evidence via meta-analysis of published studies only might produce misleading results as the published set of data may not be a representative sample of the overall evidence.<sup>1</sup> It has also been found that publicly funded research is more likely to be published irrespective of its results, whereas commercially sponsored research is more likely to be published if research findings are positive.<sup>2</sup> A thorough review of causes and examples of publication bias in medicine and psychiatry can be found elsewhere.<sup>3-5</sup> Publication bias undermines the credibility of meta-analysis results and may misinform clinical practice.

A typical example of the influence of publication bias was given by Turner *et al*<sup>6</sup> on the exaggerated efficacy of antidepressants in the published literature. The authors compared antidepressant placebo-controlled trials in the Food and Drug Administration (FDA) registry with its subset of trials that were subsequently published and found that publication bias inflated the apparent efficacy of antidepressants. Subsequently, Turner also compared antipsychotic placebo-controlled trials submitted to the FDA for regulatory approval to the related journal publications. In this case, publication bias was not so eminent probably due to their greater superiority compared with placebo.<sup>7</sup>

To overcome publication bias, mandatory registration of trials has been advocated irrespective of publication status<sup>8</sup>; however, most of the medicines prescribed today were in the market more than a decade ago and the results of old studies are not

available to the public.<sup>9</sup> Prospective registration of studies and public disclosure of their results are the only viable solutions to the publication problem, but there are still clear gaps in the trial registration system, and the risk of having biased estimates in meta-analysis is still high. Therefore, a number of data-based visual and statistical methods have been developed to explore and account for publication bias in a set of trials.

In this paper, we reviewed the most frequently used methods to assess publication bias and we used the antidepressant trials dataset by Turner to illustrate the main characteristics of this series of visual and statistical methods.

## METHODS

There are two different categories of methods for publication bias: methods based on funnel plots and selection models. Methods were applied to the FDA registered and the published trials to explore the differences between the two datasets.

### Methods based on funnel plots

#### Funnel plot

A funnel plot is a useful graph to detect small-study effects. An asymmetric scatter of studies around the summary estimate of the meta-analysis is often mistakenly equated with publication bias. However, asymmetry should not be attributed to publication bias<sup>10</sup> but should be seen as a means of examining for small-study effects.<sup>11 12</sup> Asymmetry may be caused by other reasons such as true heterogeneity in the underlying treatment effects, selective outcome reporting or chance.<sup>1</sup> Clinical heterogeneity among patients may result in an asymmetric funnel plot. An intervention may be more effective in high-risk populations which are harder to recruit, but studies involving such patients are very few. Also, studies conducted at high-risk populations require small sample size to achieve adequate levels of power.<sup>11</sup>

Lack of observed studies at certain regions of the plot that correspond to non-significant results may indicate that non-positive studies have not been published. It has been suggested to add contour lines indicating conventional milestones in levels of statistical significance (eg, <0.01, <0.05, <0.1) to the funnel plot as an aid to differentiate asymmetry due to publication bias from that due to other factors.<sup>13</sup>

#### Trim-and-fill method

The trim-and-fill is a method that attempts to identify and adjust results for publication bias.<sup>14</sup> The method starts by omitting small studies (trimming) until the plot becomes symmetrical and an adjusted summary effect is estimated from the remaining studies. Then, the funnel plot is replicated with the omitted studies replaced plus their 'missing' counterparts around the adjusted summary estimate (filling). The funnel plot becomes now symmetrical around the adjusted summary estimate. The trim-and-fill method provides a summary effect adjusted for publication bias and also estimates the number of unpublished studies. However, it makes the strong assumption that asymmetry in the funnel plot is solely caused by publication bias. The mechanism causing publication bias is unknown and we do not know whether the 'filled' studies would have been observed in the absence of publication bias. Simulation studies have shown that the method performs poorly in the presence of substantial between-study variation<sup>15 16</sup> as heterogeneity may be responsible for funnel plot asymmetry. Finally, the adjusted intervention effect is based on unobserved data and the method should account for increased uncertainty.

### Small-study meta-regression models

Apart from the visual inspection of (a)symmetry in a funnel plot, several tests have been developed to evaluate statistically if there is a dependence between intervention effect and trial size.<sup>10 17–20</sup> These tests are statistical analogues of the funnel plot and they are also referred to as tests for small-study effects. The most commonly cited test is a weighted regression of the intervention effect on its standard error (SE) with weights inversely proportional to the variance of the intervention effect (Egger's test).<sup>10</sup> In the absence of small-study effects, the constant in the weighted regression model should equal zero. Several other tests have been suggested.<sup>20–23</sup> Moreno *et al*<sup>18</sup> assume a linear relationship between intervention effect and variance (instead of SE). When the logarithm of the risk ratio or the odds ratio is used, there is an inherent correlation with its SE which perplexes the analysis. Harbord *et al*<sup>24</sup> suggested a test that reduces this correlation and Peters *et al*<sup>20</sup> suggested a test that assumes a linear relation between intervention effect and sample size weighted by a function of the sample size. Rucker *et al*<sup>25</sup> have also suggested a transformation of the intervention effect that eliminates its association to SE. The major disadvantage of regression-based methods is that, as funnel plots, they refer to the impact of small studies rather than to publication bias per se. It has been recommended that statistical tests for funnel plot asymmetry should be used with great caution<sup>12</sup> and should not be overinterpreted.<sup>11</sup>

### Selection models

A class of statistical approaches, called selection models, has been suggested to model the selection process (ie, the mechanism by which studies are selected for publication). A study selection process does not necessarily imply publication bias, as missing some of the undertaken studies might not alter the summary estimate (although it will decrease its precision).

Selection models allow researchers to evaluate the likely impact the missing studies would have, had they been included in the meta-analysis. In the selection model, we assume that the observed sample of studies is not at random; we observe the studies because they have certain characteristics that make them 'publishable' or, in other words, increase their propensity for publication. Conventional meta-analysis synthesises the data and results in a summary effect which is assumed to be unaffected by publication bias. Selection models synthesise the observed effect sizes acknowledging that the summary effect is *conditional* to the observed studies being published and identified. Then, it calculates the *marginal* effect size, which is the effect size *unconditional* to the publication status.

A selection model consists of two parts. The first part (the selection part) associates each study with an a priori probability to be published according to its features. The selection process is unknown and we resort to assumptions regarding the study characteristics associated with higher probability of publication (eg, sample size, quality of the design, etc).<sup>26</sup> The second part specifies the distribution of observed effect sizes in the published studies.

A widely used model is that suggested by Copas and Shi.<sup>27 28</sup> If published and unpublished studies do not differ in their results, the adjusted pooled estimate would be the same with that estimated from analysing the observed studies. Therefore, an important parameter to observe in a selection model (except for the summary estimate) is the correlation between observed effect size and propensity for publication. If this correlation is found to be zero, then there is no impact of the selection

process on the intervention effects. If the correlation is positive, bias arises since a large effect size entails a larger propensity for publication and the opposite for a negative correlation. In the presence of publication bias, we expect a positive correlation for beneficial outcomes and a negative for harmful outcomes.

The output of the selection model depends on the starting assumptions about the severity of selection as this is conveyed by the probabilities for large and small studies to be published. Therefore, the model can be considered under several scenarios in a sensitivity analysis.<sup>28</sup> Alternatively, we may use expert opinion to inform the probabilities of publication, their values as well as other factors that they might depend on.<sup>29</sup> We used expert opinion in the analysis of the antidepressant trials.

## RESULTS

Turner identified 73 trials (74 originally but two were subsequently combined) registered with the FDA comparing 12 antidepressants to placebo. From the 38 studies in the FDA registry with statistically significant results, only one was not published, whereas from the 36 FDA with non-statistically significant results, only three were published and another 11 were published with results conflicting those presented in the FDA report.<sup>6</sup> The summary estimate (standardised mean difference, SMD) from the published trials is 0.41 (95% CI 0.37 to 0.45). The synthesis of trials in the FDA registry yielded a summary estimate of 0.31 (95% CI 0.27 to 0.35). Hence, analysis of published studies exaggerates efficacy by 33%.

### Methods based on funnel plots

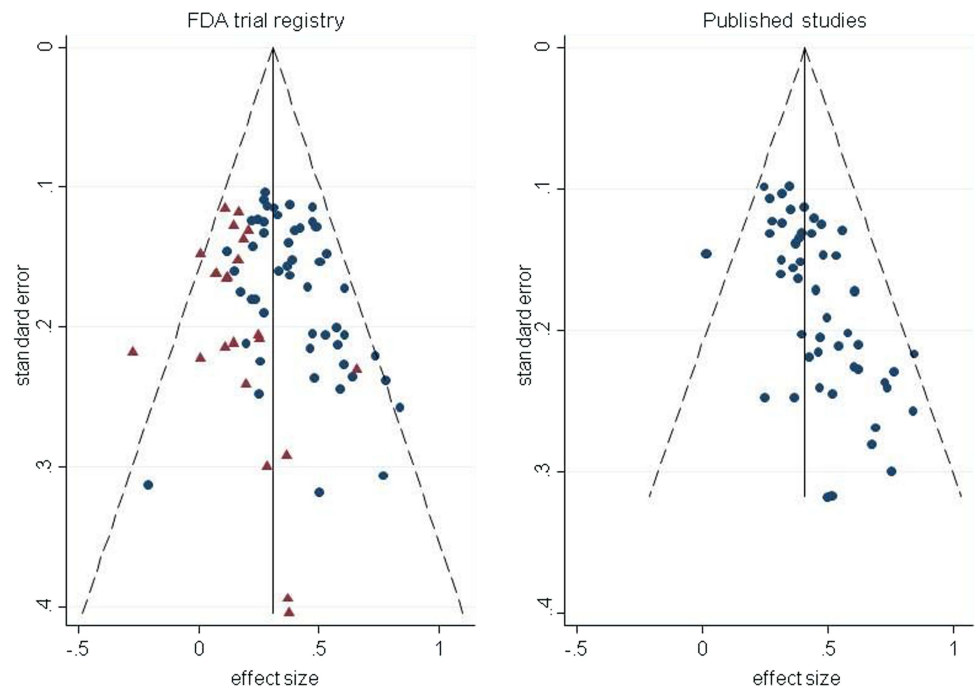
#### Funnel plot

Figure 1 shows the funnel plots for the 73 trials registered with the FDA (left-hand side plot) and the 50 published studies (right-hand side plot). The funnel plot for the studies in the FDA registry appears to be scattered symmetrically around the summary estimate which equals 0.31. Triangles refer to studies that were not subsequently published. Taking aside the unpublished studies (triangles), there is a clear pattern in the funnel plot with smaller studies showing larger effect sizes. The funnel plot that refers to the published studies is clearly asymmetrical with a trend showing that smaller studies are associated with larger effects. The summary estimate has increased to 0.41 and we see that most of the small studies show a large effect.

Figure 2 shows the contour-enhanced funnel plots for the two sets of trials. The plots allow us to visualise if there are non-positive studies in our dataset (studies lying in the large white area around zero). The left-hand side plot that refers to the FDA trial registry shows that almost half of the studies are non-positive and the proportion of such studies is the same for large and small studies. This adds credence to the possibility that a relationship between SE and effect size (asymmetry) is due to reasons other than publication bias. The right-hand side plot that refers to the published studies shows that the vast majority of studies are statistically significant favouring the antidepressants. It is also clear that studies appear to be missing in areas of statistical non-significance. It is clear that the larger the effect size in a trial the larger is its probability of publication.

Studying the likely impact of publication bias on the apparent efficacy of interventions is difficult in practice, as it requires assumptions to be made about the outcome in studies that are not actually observed. An overcautious strategy is not to pool

**Figure 1** Funnel plots with pseudo 95% confidence limits. The left-hand side plot showing studies registered with the Food and Drug Administration (FDA). Triangles refer to studies that were not subsequently published and circles to studies that were subsequently published in a scientific journal. The right-hand side plot refers to studies published in scientific journals only.



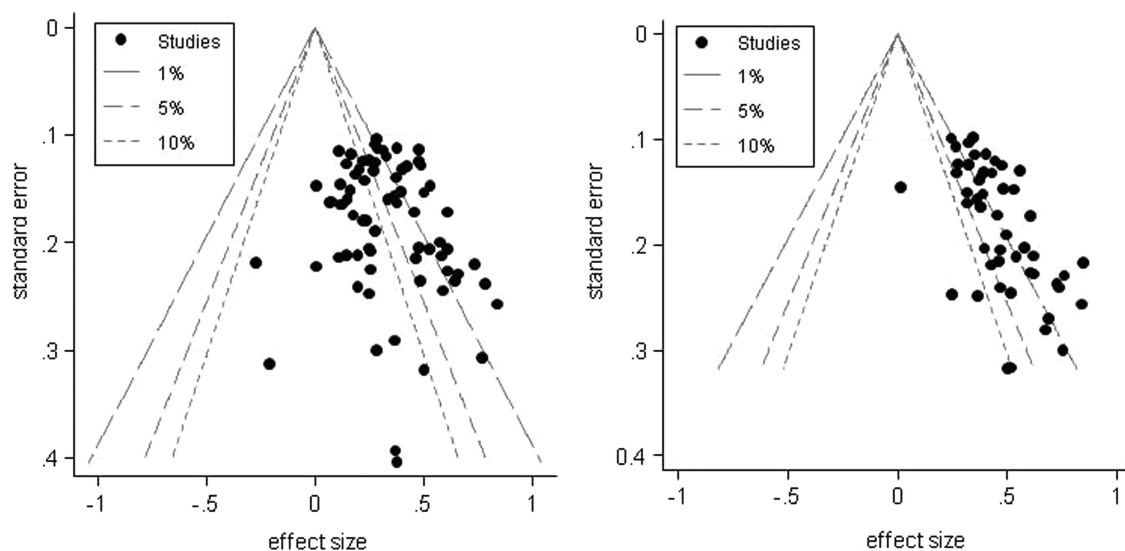
studies. Less radical strategies have been developed to adjust the summary estimate for the possible presence of publication bias.

**Trim-and-fill method**

Figure 3 shows the trimmed and filled funnel plot. The left-hand side plot that refers to the FDA trial registry shows (in squares) that eight studies were added to produce a symmetric funnel plot around an adjusted summary estimate of 0.29 (dashed line). The summary estimate is almost identical to the estimated one from the FDA trial registry (0.31, solid line). The right-hand side plot that refers to the published studies suggests that 18 studies need to be filled to create a symmetric funnel plot around an adjusted summary estimate. The summary estimate has now decreased from 0.41(95% CI 0.37 to 0.45) to 0.36 (95% CI 0.31 to 0.40).

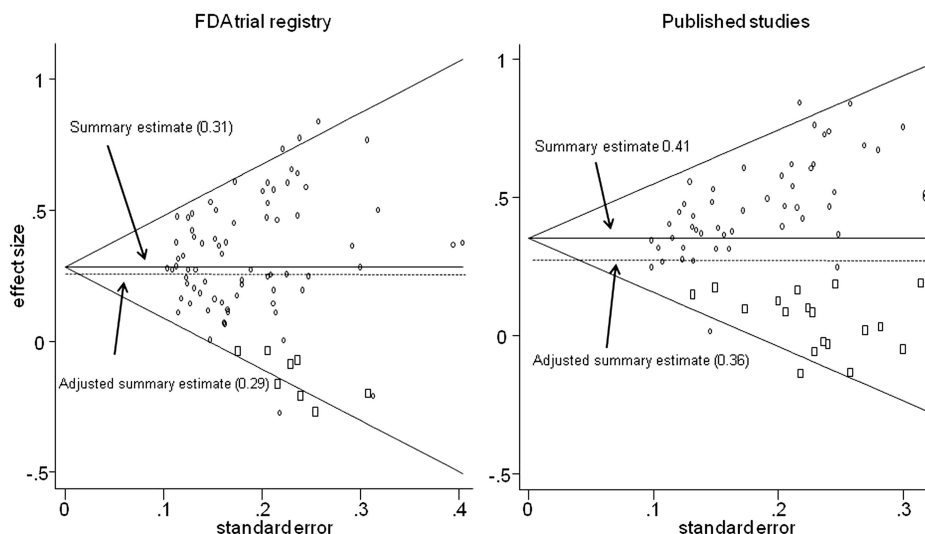
**Small-study meta-regression models**

Figure 4 shows the results of Egger’s test for the two sets of trials. An association is found when the 95% CI for the constant in the regression model (vertical line in the beginning of each plot) does not include zero. The left-hand side plot refers to the FDA trial registry while the right-hand side refers to the published studies. In the FDA registry, Egger’s test did not find a significant association between intervention effect and SE, whereas it found an association in the published literature. The hypothesis being tested is that there are no small-study effects or alternatively that the funnel plot is symmetric. The p value corresponding to this hypothesis is 0.10 in the FDA trial registry (suggesting that there are no small-study effects) whereas it is zero in the published studies. Another well-known test for determining an association between effect size and SE is the



**Figure 2** Contour-enhanced funnel plots. The left-hand side plot showing studies registered with the Food and Drug Administration (FDA). The right-hand side plot refers to studies published in scientific journals.

**Figure 3** Filled funnel plots with pseudo 95% confidence limits. The left-hand side plot showing studies registered with the Food and Drug Administration (FDA). The right-hand side plot refers to studies published in scientific journals.



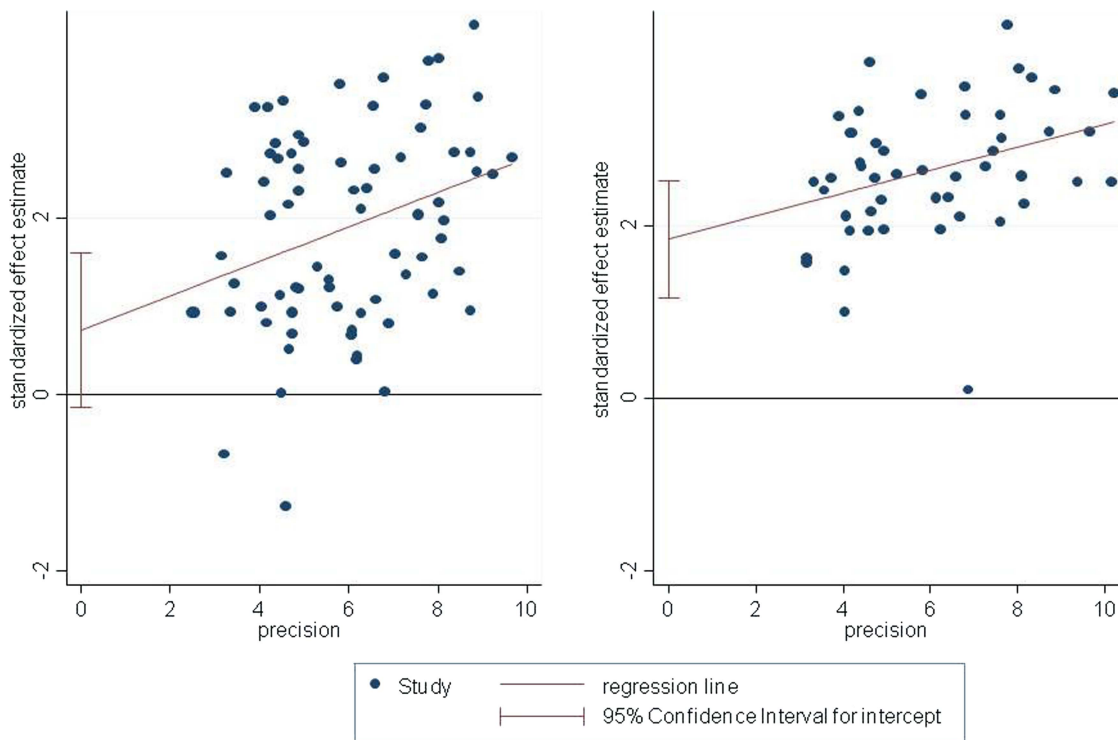
rank correlation test.<sup>21</sup> The rank correlation test gave a zero p value for the published studies and a p value of 0.07 (marginally non-significant) for the FDA dataset.

Regression-based tests have been found to perform well unless there are a few studies in the meta-analysis which is often the case.<sup>19</sup> If funnel plot asymmetry is due to bias rather than heterogeneity, results from larger studies are more trustworthy. Extrapolating a regression line on a funnel plot to a study of infinitely large sample size produces an estimate that can be regarded as ‘adjusted’ for small-study effects. In our example, extrapolating Egger’s regression line in the published dataset gave an adjusted estimate of SMD 0.13 (95% CI 0.03 to

0.24), whereas in the FDA-registry the adjusted estimate is 0.20 (95% CI 0.06 to 0.34). These contradictory results are difficult to interpret. The regression coefficient in the published studies is larger and has more discriminating power and gives a small intervention effect for very large studies.

**Selection models**

We used expert opinion to inform the selection process. More specifically, we asked nine psychiatrists, what is the probability of publication for small and large studies that compare an anti-depressant to placebo and we considered the average in their response. The experts gave a probability of publication ranging



**Figure 4** Egger’s method to test asymmetry of the funnel plot. The left-hand side plot showing studies registered with the Food and Drug Administration (FDA). The right-hand side plot refers to studies published in scientific journals. The slope of the regression line corresponds to the summary effect if there is no bias. The test for small-study effects is taken by inspecting the 95% CI for the intercept (vertical line in the beginning of each plot). If this line crosses zero, there are no small-study effects.

from 35% to 45% for a study with 40 patients per arm. The probability interval rises to 75–85% for a study with 400 patients per arm. We applied the selection model on the published studies and achieved an estimate of 0.35 (95% CI 0.31 to 0.39). The 95% CI does not include the estimated summary estimate from the published studies (without making any adjustments for publication bias) which was estimated to be 0.41. The correlation between effect size and propensity for publication is estimated to be 0.81 (95% CI 0.52 to 0.99). This means that the observed effect sizes are highly correlated with a large probability of publication. Note that this association is not confounded by heterogeneity as in the case of the meta-regression and funnel-plot approach.

We applied the selection model to the trials in the FDA registry and the unconditional summary estimate was estimated as 0.28 (95% CI 0.22 to 0.34). The 95% CI includes the summary estimate from the FDA-registered trials (0.31), suggesting that the FDA registry is an unbiased database. The correlation between study treatment effects and propensity for publication was 0.30 (95% CI -0.11 to 0.64); the CI included zero, suggesting that the magnitude of the intervention effect was not correlated with the probability of trial publication (no publication bias).

The propensity of publication was associated with the magnitude of the intervention effect and its SE. We could have added other proxies such as the quality of the study or funding information.

## DISCUSSION

A variety of methods exist for detecting and accounting for publication bias. However, there is still no consensus on which method to use while issues of understanding by non-technical audiences and ease of application play an important role in the uptake or not of the various methods. The starting point is usually a visual inspection of the funnel plot. Its assessment is often objective and an asymmetric funnel plot should not be mistakenly taken as a proof for publication bias. Regression-based tests have been suggested for a formal assessment of asymmetry. However, these tests often lack power and results are sensitive to the measure of precision used (eg, variance or SE). The funnel plot-related methods examine an association between study size and effect size rather than addressing publication bias. The trim-and-fill method is intuitively appealing since it relates publication bias to missing studies and asymmetry in the funnel plot. However, it assumes that publication bias is solely responsible for funnel plot asymmetry which is unrealistic. Enhancing the funnel plot with contours of statistical significance may help disentangle publication bias from other reasons of funnel plot asymmetry. However, there is no guarantee that studies with negative or null results have been conducted and not been published.

Statistical methods used to detect and account for small-study effects are related to publication bias, but they do not address the issue. There is no guarantee that small studies showing non-significant results have been conducted and remained unpublished. Selection models explore publication bias and provide adjusted estimates via a sensitivity analysis which is the natural way of handling unobserved data. Despite the fact that selection models are the only methodology that addresses publication bias, their theoretical mechanism and their application are not easily accessible to systematic reviewers.

**Dimitris Mavridis,<sup>1,2</sup> Georgia Salanti<sup>1</sup>**

<sup>1</sup>Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece

<sup>2</sup>Department of Primary Education, University of Ioannina, Ioannina, Greece

**Correspondence to:** Dr Dimitris Mavridis, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece; [dimi.mavridis@googlemail.com](mailto:dimi.mavridis@googlemail.com)

**Competing interests** DM and GS received research funding from the European Research Council (IMMA 260559).

## REFERENCES

- Higgins JPT, Green S. *Cochrane handbook for systematic reviews of interventions*. The Cochrane Collaboration, 2013.
- Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA* 1990;**263**:1385–9.
- Gilbody SM, Song F, Eastwood AJ, *et al*. The causes, consequences and detection of publication bias in psychiatry. *Acta Psychiatr Scand* 2000;**102**:241–9.
- Lee K, Bacchetti P, Sim I. Publication of clinical trials supporting successful new drug applications: a literature analysis. *PLoS Med* 2008;**5**:e191.
- Turner EH. Publication bias, with a focus on psychiatry: causes and solutions. *CNS Drugs* 2013;**27**:457–68.
- Turner EH, Matthews AM, Linardatos E, *et al*. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008;**358**:252–60.
- Turner EH, Knopfelmacher D, Shapley L. Publication bias in antipsychotic trials: an analysis of efficacy comparing the published literature to the US Food and Drug Administration database. *PLoS Med* 2012;**9**:e1001189.
- De Angelis C, Drazen JM, Frizelle FA, *et al*. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *Lancet* 2004;**364**:911–12.
- Goldacre B. Are clinical trial data shared sufficiently today? No. *BMJ* 2013;**347**:f1880.
- Egger M, Smith GD, Schneider M, *et al*. Bias in meta-analysis detected by a simple graphical test. *BMJ* 1997;**315**:629–34.
- Sterne JA, Egger M, Smith GD. Systematic reviews in health care: investigating and dealing with publication and other biases in meta-analysis. *BMJ* 2001;**323**:101–5.
- Sterne JA, Sutton AJ, Ioannidis JP, *et al*. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;**343**:d4002.
- Peters JL, Sutton AJ, Jones DR, *et al*. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J Clin Epidemiol* 2008;**61**:991–6.
- Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 2000;**56**:455–63.
- Peters JL, Sutton AJ, Jones DR, *et al*. Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Stat Med* 2007;**26**:4544–62.
- Terrin N, Schmid CH, Lau J, *et al*. Adjusting for publication bias in the presence of heterogeneity. *Stat Med* 2003;**22**:2113–26.
- Moreno SG, Sutton AJ, Turner EH, *et al*. Novel methods to deal with publication biases: secondary analysis of antidepressant trials in the FDA trial registry database and related journal publications. *BMJ* 2009;**339**:b2981.
- Moreno SG, Sutton AJ, Ades AE, *et al*. Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BMC Med Res Methodol* 2009;**9**:2.
- Moreno SG, Sutton AJ, Ades AE, *et al*. Adjusting for publication biases across similar interventions performed well when compared with gold standard data. *J Clin Epidemiol* 2011;**64**:1230–41.
- Peters JL, Sutton AJ, Jones DR, *et al*. Comparison of two methods to detect publication bias in meta-analysis. *JAMA* 2006;**295**:676–80.
- Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994;**50**:1088–101.
- Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Stat Med* 2006;**25**:3443–57.
- Macaskill P, Walter SD, Irwig L. A comparison of methods to detect publication bias in meta-analysis. *Stat Med* 2001;**20**:641–54.
- Harbord RM, Harris RJ, Sterne JAC. Updated tests for small-study effects in meta-analyses. *Stata J* 2009;**9**:197–210.
- Rucker G, Schwarzer G, Carpenter J. Arcsine test for publication bias in meta-analyses with binary outcomes. *Stat Med* 2008;**27**:746–63.
- Sutton AJ, Song F, Gilbody SM, *et al*. Modeling publication bias in meta-analysis: a review. *Stat Methods Med Res* 2000;**9**:421–45.
- Copas J, Shi JQ. Meta-analysis, funnel plots and sensitivity analysis. *Biostatistics* 2000;**1**:247–62.
- Copas JB, Shi JQ. A sensitivity analysis for publication bias in systematic reviews. *Stat Methods Med Res* 2001;**10**:251–65.
- Mavridis D, Sutton A, Cipriani A, *et al*. A fully Bayesian application of the Copas selection model for publication bias extended to network meta-analysis. *Stat Med* 2013;**32**:51–66.