



Page, M. J., Sterne, J. A. C., Higgins, J. P. T., & Egger, M. (2020). Investigating and dealing with publication bias and other reporting biases in meta-analyses: a review. *Research Synthesis Methods*. <https://doi.org/10.1002/jrsm.1468>

Peer reviewed version

Link to published version (if available):  
[10.1002/jrsm.1468](https://doi.org/10.1002/jrsm.1468)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Wiley at <https://doi.org/10.1002/jrsm.1468>. Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# **Investigating and dealing with publication bias and other reporting biases in meta-analyses of health research: a review**

Matthew J Page<sup>1</sup>, Jonathan AC Sterne<sup>2,3</sup>, Julian PT Higgins<sup>2</sup>, Matthias Egger<sup>4</sup>

1. School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia
2. Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
3. NIHR Bristol Biomedical Research Centre, University of Bristol, Bristol, UK
4. Institute for Social and Preventive Medicine, University of Bern, Bern, Switzerland

**Correspondence to:** Dr. Matthew Page, School of Public Health and Preventive Medicine, Monash University, 553 St Kilda Road, Melbourne, Victoria, 3004, Australia. Telephone: +61 9903 0248. Email address: [matthew.page@monash.edu](mailto:matthew.page@monash.edu)

**Word count:** 4,143 main text, 184 abstract

## **ABSTRACT**

A P value, or the magnitude or direction of results can influence decisions about whether, when, and how research findings are disseminated. Regardless of whether an entire study or a particular study result is unavailable because investigators considered the results to be unfavourable, bias in a meta-analysis may occur when available results differ systematically from missing results. In this paper, we summarize the empirical evidence for various reporting biases that lead to study results being unavailable for inclusion in systematic reviews, with a focus on health research. These biases include publication bias and selective nonreporting bias. We describe processes that systematic reviewers can use to minimize the risk of bias due to missing results in meta-analyses of health research, such as comprehensive searches and prospective approaches to meta-analysis. We also outline methods that have been designed for assessing risk of bias due to missing results in meta-analyses of health research, including using tools to assess selective nonreporting of results, ascertaining qualitative signals that suggest not all studies were identified, and generating funnel plots to identify small-study effects, one cause of which is reporting bias.

**Keywords:** Publication bias; Systematic review; Meta-analysis; Reporting

## 1. INTRODUCTION

Consider the (fictional) city of Melstol. The council recently called residents to vote on a proposal to ban cheering and clapping at the local football stadium to placate noise-sensitive residents. Passionate campaigning on both sides of the debate led to a record turnout on election day. Officials declared a close victory for the ban, and supporters rejoiced noiselessly with vigorous air punches. However, journalists later discovered that officials had withheld voting forms for 10% of the electorate living in an area that overwhelmingly opposed the measure. A recount including the suppressed votes overturned the original result, to the relief of diehard football fans. Yet all residents remain concerned by the systematic suppression of votes and credibility of the council was dented.

Systematic reviewers seeking to identify all relevant evidence face a similar situation. Study investigators may make decisions about dissemination of their research findings based on P values, or the magnitude or direction of their results. Results that are not available to reviewers may therefore differ systematically from those that are. The phenomenon is widely known as reporting bias, although it might be described more accurately as nonreporting bias<sup>1,2</sup>. Omission of relevant study results can bias the results of a meta-analysis, putting the credibility of the review in doubt. Reporting biases can also lead to bias in published results, if they are selected for publication from multiple analyses of the same association (described as bias in selection of the reported result<sup>3,4</sup>).

This paper is a modified version of a chapter from the upcoming textbook *Systematic Reviews in Health Research*<sup>5</sup>. We summarize the empirical evidence for various reporting biases that lead to study results being unavailable for inclusion in systematic reviews, with a focus on health research. We describe processes that systematic reviewers can use to minimize the risk

of bias due to missing results in a meta-analysis. We also outline different tools, plots, and statistical methods that have been designed for assessing risk of bias due to missing results in meta-analyses.

### **1.1. The evidence base for reporting biases in health research**

Basing decisions about publication on P values, or the magnitude or direction of results has traditionally been referred to as *publication bias*. To study this phenomenon, investigators have drawn samples of clinical studies from research ethics committee listings<sup>6</sup>, conference proceedings<sup>7</sup>, and regulatory submissions<sup>8</sup>, recorded which studies were published in journal articles, and examined the nature of the results in both the published and unpublished studies. These investigations have found that, on average, studies with statistically significant or “positive” results are more likely to be published than null or “negative” studies (Figure 1). Such an association has been observed for randomized and non-randomized studies of interventions<sup>6</sup>, diagnostic test accuracy studies<sup>9</sup>, prognostic accuracy studies<sup>10</sup> and qualitative studies<sup>11</sup>. Published randomized trials of health interventions also tend to have larger intervention effect estimates on average than unpublished trials<sup>12</sup>, which suggests that studies with smaller effects might be considered less worthy of publication.

Research can be disseminated selectively in other ways. Compared with studies with null or negative results, studies with positive results are more likely to be published earlier (time-lag bias)<sup>13-17</sup>, reported in multiple journal articles (duplicate or multiple publication bias)<sup>2,18</sup>, and cited more frequently by others (citation bias)<sup>19</sup>, although the magnitudes of the associations vary across clinical areas<sup>13</sup>. When studying pairs of randomized trial reports written by the same authors with one report published in German and the other in English, in 1997 Egger et al. found that the authors were more likely to publish trials in an English-language journal if

the results were statistically significant (language bias)<sup>20</sup>. This led to concerns that meta-analyses restricted to studies in English could exaggerate an intervention effect. However, later studies comparing trials published in English with trials published in a language other than English found the opposite direction of bias, with treatment effects slightly smaller in trials published in English than in trials published in another language<sup>12,21,22</sup>. A consequence of all these selective dissemination practices (e.g. publication bias, time-lag bias, language bias) is that the subset of studies that are included in systematic reviews may have results that are systematically different from studies that are less readily accessible.

Even when a study report is available, results for some outcomes that were assessed may be missing or incompletely reported because of the P value, or the magnitude or direction of the results (*selective nonreporting bias* or *outcome reporting bias*). Several studies have compared journal article reports of studies with their corresponding protocols<sup>23,24</sup>, trials register entries<sup>25,26</sup>, or documents submitted to regulators<sup>27</sup>, and identified frequent nonreporting of results for outcomes that were prespecified. The largest such study compared 227 protocols and amendments with 333 matching articles published between 1990 and 2008<sup>24</sup> and found that 7% of protocol-defined primary outcomes and 19% of secondary outcomes were not reported. In a more recent analysis, the COMPare study found that results were missing for 42% of outcomes prespecified in 67 trials published between October 2015 and January 2016 in the world's top five general medical journals<sup>28</sup>. Other studies suggest that statistically significant results for beneficial outcomes had higher odds of being completely reported than nonsignificant results<sup>29</sup> (Figure 2).

## **2. APPROACHES TO MINIMIZE RISK OF BIAS DUE TO MISSING RESULTS**

Regardless of whether an entire study or a subset of its results is unavailable, the potential consequence for a meta-analysis is bias. In this section, we describe two strategies systematic reviewers can use to minimize the risk of bias due to missing results in a meta-analysis: searching beyond journal articles, and restricting meta-analyses to inception cohorts.

## **2.1 Searching beyond journal articles**

Results of health research are available in various sources other than bibliographic databases of published studies like MEDLINE. No single resource gives access to all studies, so systematic reviewers should strive to search multiple sources. These include study registers such as ClinicalTrials.gov, databases of conference proceedings, dissertations and other “grey literature” (e.g. unpublished government and institutional reports), and grants databases such as NIH RePORTER<sup>30</sup>. Also, contacting authors or sponsors of studies may yield additional information. For clinical trials of regulated interventions (drugs and devices), reviewers should also consider consulting manufacturer registers such as the GlaxoSmithKline Study Register, websites of regulatory agencies like the European Medicines Agency, and health technology assessment agencies such as the International Network of Agencies for Health Technology Assessment<sup>30,31</sup>.

Many studies have shown that including results from sources other than journal articles can influence the magnitude or precision of meta-analytic effects<sup>32-35</sup>. For example, Mayo-Wilson et al. found that depending on whether the study data source was a journal article, conference abstract, regulatory document, or individual participant dataset, they were able to produce a meta-analytic result showing that an agent was effective for pain relief in some instances and ineffective in others<sup>36</sup>, which shows how valuable it can be to search beyond standard bibliographic databases.

Identifying and using data in sources other than journal articles can present various challenges. For example, search interfaces for trials registers are relatively unsophisticated<sup>31</sup>, and there may be long delays between request and receipt of data from regulators or manufacturers<sup>37</sup>. Nevertheless, uncovering such data should be undertaken especially when the published studies do not report data on key outcomes that are likely to have been measured.

## **2.2 Restricting meta-analyses to inception cohorts**

Study identification is retrospective in most systematic reviews; that is, authors typically search for reports of *completed* studies. However, given the evidence that completed studies with positive results are easier to identify because of the nature of their results, systematic reviews based on completed studies are likely to include a biased subset of all studies conducted. To minimize biased inclusion of studies, systematic reviewers might instead synthesize results of studies that could be identified before their results became known (an inception cohort). For example, different teams of researchers could work together to design and conduct a set of studies addressing the same question, and synthesize the results once all studies are completed (prospective meta-analysis)<sup>38</sup>. Assuming the researchers agreed to include all relevant results in the analysis, there would be no risk of bias due to missing results in the prospective meta-analysis. Few prospective meta-analyses have been conducted, but numbers are increasing: a systematic search for all prospective meta-analyses published up to February 2018 identified 43, almost half of which were published from 2015 onwards<sup>39</sup>.

Prospective meta-analysis is recommended for high priority research questions for which limited previous evidence exists and new studies are expected to emerge, and is not suitable in all cases<sup>39</sup>. However, systematic reviewers can generate an inception cohort in other ways, for



example, by restricting inclusion to prospectively registered (or pre-registered) studies addressing the review question<sup>40</sup>, or identifying trials from registries before results were generated and working together with the trialists to populate the meta-analysis<sup>41</sup>. If all relevant results are available for all prospectively registered studies, there will be no bias due to missing results in a meta-analysis of these studies. If results are unavailable for some of the prospectively registered studies, then authors using this approach will be able to quantify how much evidence is missing, unlike a standard (retrospective) systematic review. However, a limitation of this approach is that the precision of a meta-analysis may be low if there are only a few, small, prospectively registered studies addressing the review question. Restricting a synthesis to an inception cohort may therefore involve a trade-off between bias and precision.

### **3 APPROACHES TO ASSESS RISK OF BIAS DUE TO MISSING RESULTS**

Researchers have developed many approaches seeking to assess selective publication or reporting of study results, and the impact this may have on a meta-analysis. A systematic search for scales and checklists designed to help authors make a qualitative judgement about the risk of reporting biases identified 15 tools published up to February 2017<sup>42</sup>. The tools varied by the type of reporting bias (publication bias or selective nonreporting bias) assessed; the target of assessment (e.g. an individual study or a meta-analysis of studies); and the criteria used to designate a study or meta-analysis as at risk of bias. A systematic search for graphical and statistical approaches designed to detect or adjust for reporting biases identified nearly 100 methods published up to January 2013<sup>43</sup>, and additional methods have been developed since<sup>44</sup>. However, all these approaches have limitations, and few have been validated empirically using examples in which the true amount of missing evidence was known. In the following section, we provide an overview of some of the available methods.

### 3.1 Tools to assess selective nonreporting of results in the identified studies

Various tools have been developed to assess selective nonreporting of results<sup>42</sup>. All emphasize the importance of retrieving the study's protocol, registration record, or statistical analysis plan so that the planned outcomes and analyses can be compared with those that were reported<sup>42</sup>. If study plans are not available (which is more likely to be the case for older studies, and for non-randomized studies), an assessment of selective non-reporting is still possible. For example, review authors can check whether any outcomes listed in the Methods section of a report are incompletely reported or have no corresponding results available in the Results section. By "incompletely reported" we mean that the study authors present insufficient data for inclusion of the result in a meta-analysis (for example, stating only that the between-group difference in the number of deaths was not significant, rather than reporting the number of deaths in each group or the risk ratio and 95% confidence interval). Regardless of the study design, users can also gauge the likelihood that a particular outcome of interest was measured, taking into consideration factors such as the clinical importance of the outcome. For example, pain is a defining symptom of shoulder disorders<sup>45</sup>, so its absence in a trial report may raise suspicion of selective nonreporting.

An approach commonly used to assess selective nonreporting of results was via one of the domains of the 2011 Cochrane risk of bias tool for randomized trials<sup>46</sup>. The tool asks users to judge the risk of selective nonreporting bias in a *study* as either low, unclear, or high, and to provide reasons for their judgement. This approach has limitations<sup>47,48</sup>. Study-level assessments inform readers which studies the systematic reviewers have concerns about, but not necessarily which results were incompletely reported or missing entirely from those studies. An audit of Cochrane reviews published in 2015 found that in 39% of studies rated at high risk of bias due to selective nonreporting, users of the risk of bias tool failed to specify the particular

results that were incompletely reported<sup>48</sup>. Outcome-level assessments, as recommended by the ORBIT (Outcome Reporting Bias In Trials) tool<sup>49</sup>, can overcome this problem, by displaying which results are unavailable for which studies, and whether the reasons for unavailability give cause for concern.

A limitation of existing tools for assessing selective nonreporting of results is that they do not guide reviewers to assess the risk of bias in meta-analyses that are unable to include the selectively nonreported results. This may explain why only 30% of Cochrane review authors who declared suspicion of selective nonreporting in their included studies acknowledged that the meta-analyses presented in the review were missing results<sup>48</sup>. A new framework that addresses these problems has recently been developed<sup>50</sup>.

### **3.2 Qualitative signals for additional missing results**

Some tools for assessing risk of reporting biases guide users to consider various qualitative signals that suggest additional results may be missing from studies that have not been identified<sup>42</sup>. These signals include:

- Sources of unpublished studies (e.g. trials registers) were not consulted.
- Specialized bibliographic databases that are likely to index studies relevant to the review question were not consulted.
- Only English-language studies were eligible but the review addresses a question frequently investigated in countries speaking a language other than English.
- The research area addressed by the review is fast-moving (hence there is a risk of time-lag bias).
- There is prior evidence of reporting bias in the research area addressed by the review.

The presence of one or more of these signals does not prove that additional results are missing from a particular meta-analysis. However, considering them is useful when trying to reach an overall judgement about risk of bias due to missing results, particularly in cases where information on prespecified outcomes and analyses is unavailable for most studies.

### **3.3 Funnel plots**

Funnel plots have long been used to assess the possibility that results are systematically missing from a meta-analysis. However, they should not be considered to be diagnostic of the presence of reporting biases because several other factors influence their appearance<sup>51</sup>. In this section, we describe what funnel plots are and how to interpret them.

First used in educational research and psychology<sup>52</sup>, a funnel plot is a simple scatter plot of the intervention effects estimated from individual studies on the x axis against some measure of study size on the y axis, typically the standard error of the effect estimate<sup>53</sup>. The name “funnel plot” is based on the fact that the precision of the estimate of the underlying intervention effect will increase as study sample size increases; effect estimates from small studies will therefore scatter more widely at the bottom of the graph, with the spread narrowing among larger studies. The plot will resemble a symmetric, inverted funnel if there is no bias or between-study heterogeneity, and hence the scatter is due to sampling variation alone (see panel A of Figure 3).

Reporting biases are one of several factors that may lead to asymmetry in a funnel plot (Table 1). For example, if smaller studies showing no statistically significant effects remain unpublished, then such publication bias will lead to an asymmetrical appearance of the funnel plot with a gap in the bottom corner of the graph (see panel B of Figure 3). However, studies

with less methodological rigour tend to show larger intervention effects<sup>54</sup>, so asymmetry also can arise when some smaller studies are at higher risk of bias and therefore produce larger intervention effect estimates (see panel C of Figure 3). Therefore, the funnel plot should be seen as a generic means of examining *small study effects*—the tendency for the smaller studies in a meta-analysis to show larger treatment effects—rather than a tool to diagnose specific types of bias<sup>51,55</sup>.

The studies displayed in a funnel plot may not always estimate the same underlying effect of the same intervention, and such heterogeneity between results may lead to asymmetry in funnel plots if the true intervention effect is larger in the smaller studies. For example, randomized trials conducted in high risk patients will tend to be smaller because of the difficulty in recruiting such patients, and because increased event rates mean that smaller sample sizes are required to detect a given effect<sup>56</sup>. Small trials generally are conducted before larger trials are established, and in the intervening years standard, control treatments may have improved, which can reduce the relative efficacy of the experimental treatment. Trialists may have implemented interventions less thoroughly in larger trials, thus explaining the more positive results in smaller trials<sup>57</sup>.

Some effect estimates, such as log odds ratios, are naturally correlated with their standard errors. Because of this, a funnel plot that shows no asymmetry when plotted using one effect measure could be asymmetric when plotted using a different one<sup>55</sup>. Finally, it is possible that an asymmetric funnel plot arises merely by chance.

### **3.4 Contour-enhanced funnel plots**

An enhancement to the funnel plot includes contour lines corresponding to levels of statistical significance:  $P = 0.01, 0.05, 0.1$ , etc.<sup>58</sup>. This facilitates inspection of the statistical significance of study effect estimates and whether areas in which studies seem to be missing are related to  $P$  values. Such contour-enhanced funnel plots may help systematic reviewers differentiate asymmetry due to reporting biases from bias due to the other factors described in Section 3.3.

Consider the funnel plot in Figure 4, which represents a meta-analysis of the effect of selective serotonin reuptake inhibitors (SSRIs) versus placebo on treatment response, where a risk ratio greater than 1 indicates benefit of SSRIs<sup>59</sup>. There is a suggestion of missing results in the left-hand side of the plot, where results would be unfavourable to SSRIs, and in the area of statistical nonsignificance, which adds credence to the possibility that the asymmetry is due to reporting biases. Contrast this with the funnel plot in Figure 5, which corresponds to meta-analysis of the effect of higher versus lower intake of long-chain omega-3 fats on all-cause mortality in which a risk ratio lower than 1 indicates benefit of higher intake<sup>60</sup>. In this case, there is a suggestion of missing results in the right-hand side of the plot, where results would be favourable to lower intake of omega-3 and in the area of statistical nonsignificance. However, given that almost all results in the plot appear in the area of statistical nonsignificance, i.e. both large and small studies finding a beneficial effect of higher or lower intake were nearly all not statistically significant, this reduces the plausibility that reporting bias is the underlying cause of this funnel plot asymmetry.

Funnel plot asymmetry thus may raise the possibility of bias due to missing results, but is not proof of bias. A further concern is that visual interpretation of funnel plots is inherently subjective.

### 3.5 Tests for funnel plot asymmetry

Several statistical tests for funnel plot asymmetry are available. These examine whether the association between estimated intervention effects and a measure of study size or precision is greater than that expected to occur by chance. However, after reviewing the results of simulation studies evaluating test characteristics, and based on theoretical considerations, Sterne et al. advised that tests for funnel plot asymmetry are applicable only in the minority of meta-analyses for which their use is appropriate<sup>55</sup>. For meta-analyses of randomized trials, they suggested that as a rule of thumb, tests for funnel plot asymmetry should be used only when at least 10 trials are included in the meta-analysis. This is because when there are fewer than 10 trials the power of the tests is low. If there is substantial heterogeneity, the minimum number of trials may be substantially more than 10. Sterne et al. also suggested that results of tests for funnel plot asymmetry should be interpreted in the light of visual inspection of the funnel plot. For example, when there is evidence for small study effects based on the result of an asymmetry test, it may be reasonable to exclude reporting biases as an explanation if there were very few studies with statistically significant results and bias would be expected to favour studies with statistically significant results. These recommendations apply only to meta-analyses of randomized trials, as the performance of tests for funnel plot asymmetry in other contexts (e.g. meta-analyses of prevalence, prognosis and diagnostic test accuracy studies) is likely to differ<sup>55</sup>.

Sterne et al. provided detailed recommendations about which tests to use for meta-analyses of randomized trials of intervention effects measured as mean differences, standardized mean differences, odds ratios, risk ratios, and risk differences<sup>55</sup>. Some tests, including the original and widely used Egger test<sup>51</sup>, and Begg and Mazumdar test<sup>61</sup>, are not recommended for application to odds ratios and standardized mean differences because of artefactual correlations

between the effect size and its standard error<sup>55,62,63</sup>. For odds ratios, methods proposed by Harbord et al.<sup>64</sup> and Peters et al.<sup>65</sup> overcome this problem; for standardized mean differences, see methods proposed by Zwetsloot et al.<sup>62</sup> and Pustejovsky et al.<sup>63</sup>. For tests for use in meta-analyses of trials of survival data, see Debray et al.<sup>66</sup>.

When a test for funnel plot asymmetry provides evidence of small study effects, reporting biases should be considered as one of several possible explanations (described in Section 3.3), and systematic reviewers should attempt to distinguish the different possible reasons for it. Further information on tests for funnel plot asymmetry can be found in the historical review by Marks-Anglin and Chen<sup>44</sup>.

### **3.6 Sensitivity analyses**

Statisticians have proposed several statistical approaches to assess how robust meta-analyses are to various assumptions about the extent and nature of missing results, including trim-and-fill<sup>67</sup>, selection models<sup>68</sup>, and regression-based adjustment methods<sup>69</sup>. Nearly all are designed to assess robustness to selective publication of studies<sup>70</sup>, although some are designed to assess robustness to selective nonreporting of results<sup>71,72</sup>; none assess robustness to both sources of bias. Given it is impossible to know for certain whether reporting biases have influenced the results of a review, or by how much, these methods should be considered only as sensitivity analysis, rather than as a way of ascertaining the “true” meta-analytic effect. Simulation studies have not compared the performance of all proposed methods. However, the available evidence suggests that no single method outperforms others in all scenarios, and thus there is a danger that uncritical application can lead to inappropriate conclusions being drawn<sup>68,73,74</sup>. Vevea et al.<sup>70</sup> and Marks-Anglin and Chen<sup>44</sup> summarize the advantages and disadvantages of different methods, indicate circumstances in which each can be used, and describe software available to



implement them. Given the complexity of the methods, consultation with a statistician is recommended for their implementation.

### **3.7 Summary of approaches**

We have described several approaches systematic reviewers can use to assess the risk of reporting biases. These include comparison of prespecified analysis plans with completed reports to detect selective nonreporting of results, consideration of qualitative signals that suggest not all studies were identified, and the use of funnel plots to identify small-study effects, for which reporting bias is one of several causes. Information from approaches such as funnel plots and selection models is more difficult to interpret than from less subjective approaches such as detection of incompletely reported results in studies for which prespecified analysis plans were available. Careful thought is currently being given as to how to weigh the various pieces of information gained from each approach when reaching an overall judgement of the risk of bias due to missing results in a meta-analysis. For interim guidance, see Chapter 13 of the *Cochrane Handbook for Systematic Reviews of Interventions*<sup>50</sup>.

## **4 CONCLUSION**

The evidence that dissemination of research findings can be influenced by the nature of the findings themselves is convincing; and when available results differ systematically from missing results, meta-analyses will be biased. Systematic reviewers should comprehensively search for study reports and consult not only multiple bibliographic databases but trials registers, manufacturers, regulators, and study authors and sponsors. Unless they use prospective approaches to meta-analysis, which can eliminate the potential for bias due to missing results, reviewers should formally assess the risk of bias due to missing results in their review. Several approaches can facilitate such assessment: tools to record selective

nonreporting of results, ascertaining qualitative signals that suggest not all studies were identified, and the use of funnel plots to identify small-study effects, one cause of which is reporting bias. How to weigh diverse information about the likelihood and nature of missing results in the judgement of the risk of bias in a meta-analysis is a work in progress.

## **DECLARATIONS**

### **Funding**

MJP is supported by an Australian Research Council Discovery Early Career Researcher Award (DE200101618). JACS and JPTH are National Institute for Health Research (NIHR) Senior Investigators (NF-SI-0611-10168 and NF-SI-0617-10145, respectively). JACS and JPTH are supported by NIHR Bristol Biomedical Research Centre at University Hospitals Bristol NHS Foundation Trust and the University of Bristol. JACS and JPTH are supported by the NIHR Applied Research Collaboration West (ARC West) at University Hospitals Bristol NHS Foundation Trust. JACS and JPTH are members of the MRC Integrative Epidemiology Unit at the University of Bristol. The views expressed in this article are those of the authors and do not necessarily represent those of the NHS, the NIHR, MRC, or the Department of Health and Social Care. The funders had no role in the study design, decision to publish, or preparation of the manuscript.

### **Author Contributions**

All authors declare to meet the ICMJE conditions for authorship. MJP wrote the first draft of the article. All authors contributed to revisions of the article. All authors approved the final version of the submitted article.

**Competing interests**

No competing interests were disclosed.

## **HIGHLIGHTS**

- Bias in a meta-analysis may occur when available results differ systematically from missing results.
- Several different tools, plots, and statistical methods have been designed for assessing risk of bias due to missing results in meta-analyses. These include comparison of prespecified analysis plans with completed reports to detect selective nonreporting of results, consideration of qualitative signals that suggest not all studies were identified, and the use of funnel plots to identify small-study effects, for which reporting bias is one of several causes.
- Information from approaches such as funnel plots and selection models is more difficult to interpret than from less subjective approaches such as detection of incompletely reported results in studies for which prespecified analysis plans were available.

## REFERENCES

1. Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA*. 1990;263(10):1385-1389.
2. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet*. 1991;337(8746):867-872.
3. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898.
4. Sterne JA, Hernan MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919.
5. Tanner-Smith EE, Egger M, Higgins J. Systematic reviews in health research. *Research Synthesis Methods*. 2019;10(3):310-311.
6. Schmucker C, Schell LK, Portalupi S, et al. Extent of non-publication in cohorts of studies approved by research ethics committees or included in trial registries. *PLoS One*. 2014;9(12):e114023.
7. Scherer RW, Meerpohl JJ, Pfeifer N, Schmucker C, Schwarzer G, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database of Systematic Reviews*. 2018(11):MR000005.
8. Chan A-W, Song F, Vickers A, et al. Increasing value and reducing waste: addressing inaccessible research. *Lancet*. 2014;383(9913):257-266.
9. Treanor L, Frank RA, Cherpak LA, et al. Publication bias in diagnostic imaging: conference abstracts with positive conclusions are more likely to be published. *European Radiology*. 2020;30(5):2964-2972.
10. Vollgraff Heidweiller-Schreurs CA, Korevaar DA, Mol BWJ, et al. Publication bias may exist among prognostic accuracy studies of middle cerebral artery Doppler ultrasound. *Journal of Clinical Epidemiology*. 2019;116:1-8.

11. Petticrew M, Egan M, Thomson H, Hamilton V, Kunkler R, Roberts H. Publication bias in qualitative research: what becomes of qualitative research presented at conferences? *Journal of Epidemiology and Community Health*. 2008;62(6):552-554.
12. Dechartres A, Atal I, Riveros C, Meerpohl J, Ravaud P. Association between publication characteristics and treatment effect estimates: A meta-epidemiologic study. *Annals of Internal Medicine*. 2018;169(6):385-393.
13. Tanner-Smith EE, Polanin JR. A retrospective analysis of dissemination biases in the brief alcohol intervention literature. *Psychology of Addictive Behaviors : journal of the Society of Psychologists in Addictive Behaviors*. 2015;29(1):49-62.
14. Sune P, Sune JM, Montoro JB. Positive outcomes influence the rate and time to publication, but not the impact factor of publications of clinical trial results. *PLoS One*. 2013;8(1):e54583.
15. Song SY, Koo DH, Jung SY, Kang W, Kim EY. The significance of the trial outcome was associated with publication rate and time to publication. *Journal of Clinical Epidemiology*. 2017;84:78-84.
16. Khan NA, Singh M, Spencer HJ, Torralba KD. Randomized controlled trials of rheumatoid arthritis registered at ClinicalTrials.gov: what gets published and when. *Arthritis & Rheumatology (Hoboken, NJ)*. 2014;66(10):2664-2674.
17. Korevaar DA, van Es N, Zwinderman AH, Cohen JF, Bossuyt PM. Time to publication among completed diagnostic accuracy studies: associated with reported accuracy estimates. *BMC Medical Research Methodology*. 2016;16:68.
18. Tramer MR, Reynolds DJ, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. *BMJ*. 1997;315(7109):635-640.

19. Duyx B, Urlings MJE, Swaen GMH, Bouter LM, Zeegers MP. Scientific citations favor positive results: a systematic review and meta-analysis. *Journal of Clinical Epidemiology*. 2017;88:92-101.
20. Egger M, Zellweger-Zahner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. *Lancet*. 1997;350(9074):326-329.
21. Juni P, Holenstein F, Sterne J, Bartlett C, Egger M. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *International Journal of Epidemiology*. 2002;31(1):115-123.
22. Egger M, Juni P, Bartlett C, Holenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technology Assessment*. 2003;7(1):1-76.
23. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA*. 2004;291(20):2457-2465.
24. Redmond S, von Elm E, Blumle A, Gengler M, Gsponer T, Egger M. Cohort study of trials submitted to ethics committee identified discrepant reporting of outcomes in publications. *Journal of Clinical Epidemiology*. 2013;66(12):1367-1375.
25. Jones CW, Keil LG, Holland WC, Caughey MC, Platts-Mills TF. Comparison of registered and published outcomes in randomized controlled trials: a systematic review. *BMC Medicine*. 2015;13:282.
26. Korevaar DA, Ochodo EA, Bossuyt PM, Hooft L. Publication and reporting of test accuracy studies registered in ClinicalTrials.gov. *Clinical Chemistry*. 2014;60(4):651-659.

27. Jefferson T, Doshi P, Boutron I, et al. When to include clinical study reports and regulatory documents in systematic reviews. *BMJ Evidence-Based Medicine*. 2018;23(6):210-217.
28. Goldacre B, Drysdale H, Dale A, et al. COMPare: a prospective cohort study correcting and monitoring 58 misreported trials in real time. *Trials*. 2019;20(1):118.
29. Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. *PLoS One*. 2013;8(7):e66844.
30. Lefebvre C, Glanville J, Briscoe S, et al. Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.1 (updated September 2020). Available from <https://training.cochrane.org/cochrane-handbook-systematic-reviews-interventions>: Cochrane; 2020.
31. Isojarvi J, Wood H, Lefebvre C, Glanville J. Challenges of identifying unpublished data from clinical trials: getting the best out of clinical trials registers and other novel sources. *Research Synthesis Methods*. 2018;9(4):561-578.
32. Baudard M, Yavchitz A, Ravaud P, Perrodeau E, Boutron I. Impact of searching clinical trial registries in systematic reviews of pharmaceutical treatments: methodological systematic review and reanalysis of meta-analyses. *BMJ*. 2017;356:j448.
33. Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses. *BMJ*. 2012;344:d7202.
34. Schmucker CM, Blumle A, Schell LK, et al. Systematic review finds that study data not published in full text articles have unclear impact on meta-analyses results in medical research. *PloS One*. 2017;12(4):e0176210.



35. Golder S, Loke YK, Wright K, Norman G. Reporting of Adverse Events in Published and Unpublished Studies of Health Care Interventions: A Systematic Review. *PLoS Medicine*. 2016;13(9):e1002127.
36. Mayo-Wilson E, Li T, Fusco N, et al. Cherry-picking by trialists and meta-analysts can drive conclusions about intervention efficacy. *Journal of Clinical Epidemiology*. 2017;91:95-110.
37. Doshi P, Jefferson T. Open data 5 years on: a case series of 12 freedom of information requests for regulatory data to the European Medicines Agency. *Trials*. 2016;17(1):78.
38. Thomas J, Askie LM, Berlin JA, et al. Chapter 22: Prospective approaches to accumulating evidence. In: Higgins JPT, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.1 (updated September 2020). Available from <https://training.cochrane.org/cochrane-handbook-systematic-reviews-interventions>; Cochrane; 2020.
39. Seidler AL, Hunter KE, Cheyne S, Ghersi D, Berlin JA, Askie L. A guide to prospective meta-analysis. *BMJ*. 2019;367:l5342.
40. Roberts I, Ker K, Edwards P, Beecher D, Manno D, Sydenham E. The knowledge system underpinning healthcare is not fit for purpose and must change. *BMJ*. 2015;350:h2463.
41. Sterne JAC, Murthy S, Diaz JV, et al. Association Between Administration of Systemic Corticosteroids and Mortality Among Critically Ill Patients With COVID-19: A Meta-analysis. *JAMA*. 2020;324(13):1330-1341.
42. Page MJ, McKenzie JE, Higgins JPT. Tools for assessing risk of reporting biases in studies and syntheses of studies: a systematic review. *BMJ Open*. 2018;8:e019703.

43. Mueller KF, Meerpohl JJ, Briel M, et al. Methods for detecting, quantifying and adjusting for dissemination bias in meta-analysis are described. *Journal of Clinical Epidemiology*. 2016;80:25-33.
44. Marks-Anglin A, Chen Y. A historical review of publication bias. *Research Synthesis Methods*. 2020.
45. Page MJ, Huang H, Verhagen AP, Gagnier JJ, Buchbinder R. Outcome Reporting in Randomized Trials for Shoulder Disorders: Literature Review to Inform the Development of a Core Outcome Set. *Arthritis Care Res (Hoboken)*. 2018;70(2):252-259.
46. Higgins JPT, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928.
47. Sterne JAC. Why the Cochrane risk of bias tool should not include funding source as a standard item [editorial]. *Cochrane Database of Systematic Reviews*. 2013(12):ED000076.
48. Page MJ, Higgins JPT. Rethinking the assessment of risk of bias due to selective reporting: a cross-sectional study. *Syst Rev*. 2016;5(1):108.
49. Kirkham JJ, Altman DG, Chan AW, Gamble C, Dwan KM, Williamson PR. Outcome reporting bias in trials: a methodological approach for assessment and adjustment in systematic reviews. *BMJ*. 2018;362:k3802.
50. Page MJ, Higgins JPT, Sterne JAC. Chapter 13: Assessing risk of bias due to missing results in a synthesis. In: Higgins JPT, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.1 (updated September 2020). Available from <https://training.cochrane.org/cochrane-handbook-systematic-reviews-interventions>: Cochrane; 2020.

51. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629-634.
52. Light RJ, Pillemer DB. *Summing Up: The Science of Reviewing Research*. Cambridge (MA): Harvard University Press; 1984.
53. Sterne JAC, Egger M. Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology*. 2001;54:1046-1055.
54. Page MJ, Higgins JPT, Clayton G, Sterne JAC, Hrobjartsson A, Savovic J. Empirical evidence of study design biases in randomized trials: systematic review of meta-epidemiological studies. *PLoS One*. 2016;11:7.
55. Sterne JAC, Sutton AJ, Ioannidis JPA, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*. 2011;343:d4002.
56. Davey Smith G, Egger M. Who benefits from medical interventions? Treating low risk patients can be a high risk strategy. *BMJ*. 1994;308:72-74.
57. Stuck AE, Rubenstein LZ, Wieland D. Bias in meta-analysis detected by a simple, graphical test. Asymmetry detected in funnel plot was probably due to true heterogeneity. Letter. *BMJ*. 1998;316:469-471.
58. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology*. 2008;61(10):991-996.
59. Williams T, Hattingh CJ, Kariuki CM, et al. Pharmacotherapy for social anxiety disorder (SAnD). *Cochrane Database Syst Rev*. 2017;10:CD001206.
60. Abdelhamid AS, Brown TJ, Brainard JS, et al. Omega-3 fatty acids for the primary and secondary prevention of cardiovascular disease. *Cochrane Database Syst Rev*. 2018;11:CD003177.

61. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics*. 1994;50(4):1088-1101.
62. Zwetsloot P-P, Van Der Naald M, Sena ES, et al. Standardized mean differences cause funnel plot distortion in publication bias assessments. *eLife*. 2017;6:e24260.
63. Pustejovsky JE, Rodgers MA. Testing for funnel plot asymmetry of standardized mean differences. *Research Synthesis Methods*. 2019;10(1):57-71.
64. Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Statistics in Medicine*. 2006;25(20):3443-3457.
65. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of two methods to detect publication bias in meta-analysis. *JAMA*. 2006;295(6):676-680.
66. Debray TPA, Moons KGM, Riley RD. Detecting small-study effects and funnel plot asymmetry in meta-analysis of survival data: A comparison of new and existing tests. *Research Synthesis Methods*. 2018;9(1):41-50.
67. 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(2):455-463.
68. McShane BB, Bockenholt U, Hansen KT. Adjusting for Publication Bias in Meta-Analysis: An Evaluation of Selection Methods and Some Cautionary Notes. *Perspectives on Psychological Science*. 2016;11(5):730-749.
69. 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.
70. Vevea JL, Coburn K, Sutton A. Publication bias. In: Cooper H, Hedges LV, Valentine JC, eds. *The Handbook of Research Synthesis and Meta-Analysis*. Russell Sage Foundation; 2019:383-430.

71. Copas J, Marson A, Williamson P, Kirkham J. Model-based sensitivity analysis for outcome reporting bias in the meta analysis of benefit and harm outcomes. *Statistical Methods in Medical Research*. 2019;28(3):889-903.
72. Williamson PR, Gamble C. Application and investigation of a bound for outcome reporting bias. *Trials*. 2007;8:9.
73. Carter EC, Schönbrodt FD, Gervais WM, Hilgard J. Correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science*. 2019;2(2):115-144.
74. Rucker G, Carpenter JR, Schwarzer G. Detecting and adjusting for small-study effects in meta-analysis. *Biometrical Journal Biometrische Zeitschrift*. 2011;53(2):351-368.
75. Chan AW, Krleža-Jeric K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *Canadian Medical Association Journal*. 2004;171(7):735-740.

**Table 1. Possible sources of asymmetry in funnel plots (adapted from Egger et al.<sup>51</sup>)**

---

1. Publication bias and other reporting biases
<ul style="list-style-type: none"><li>• Entire study reports, or particular results, of smaller studies are unavailable because of the P value, magnitude or direction of effect.</li></ul>
2. Poor methodological quality leading to spuriously inflated effects in smaller studies
<ul style="list-style-type: none"><li>• Asymmetry can arise when some smaller studies are of lower methodological quality and produce larger intervention effect estimates.</li></ul>
3. True heterogeneity
<ul style="list-style-type: none"><li>• Substantial benefit may be seen only in patients at high risk for the outcome that is affected by the intervention, and usually these high-risk patients are more likely to be included in small, early studies<sup>56</sup>.</li><li>• Some interventions may have been implemented more thoroughly in smaller trials and may, therefore, have resulted in larger intervention effect estimates<sup>57</sup>.</li></ul>
4. Artefactual
<ul style="list-style-type: none"><li>• Some effect estimates are naturally correlated with their standard errors, and this can produce spurious asymmetry in a funnel plot<sup>55,62</sup></li></ul>
5. Chance

---

## Figure legends

Figure 1. Random-effects meta-analysis of meta-analyses investigating the association between publication status and the P value, magnitude or direction of the results. Data for the meta-analysis of studies submitted for research ethics committee approval comes from Schmucker et al.<sup>6</sup>. Data for the meta-analysis of randomized trials presented at conferences comes from Scherer et al.<sup>7</sup>. Data for the meta-analysis of clinical trials submitted for regulatory approval comes from Chan et al.<sup>8</sup>.

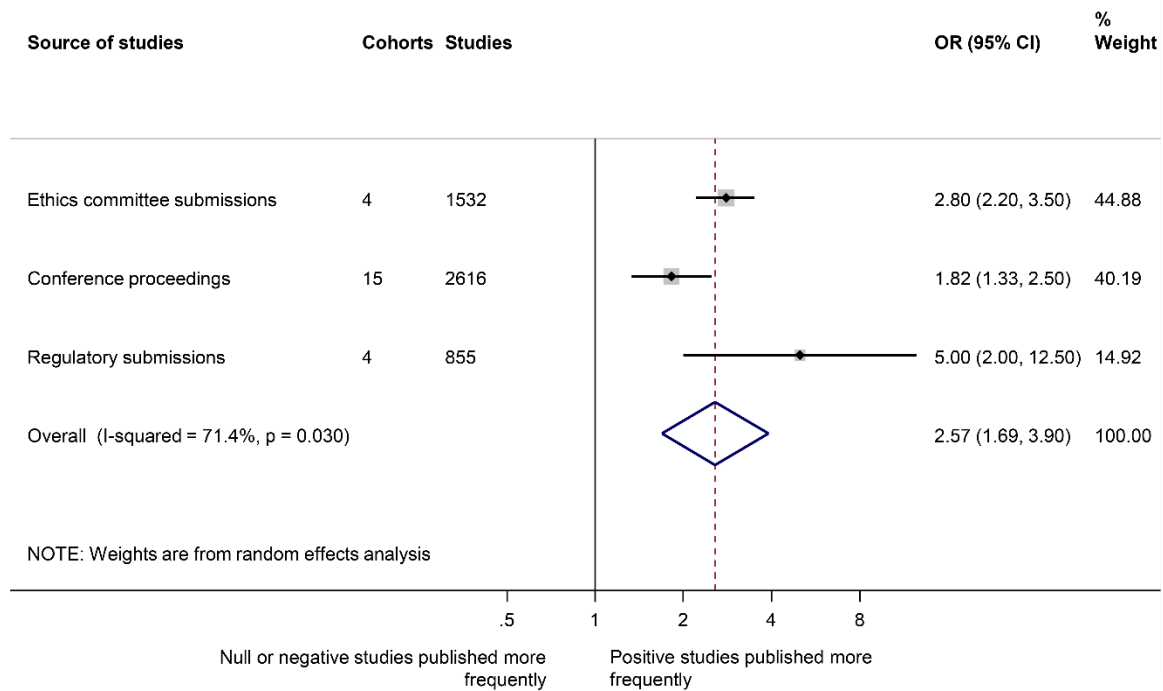


Figure 2. Random-effects meta-analysis of studies investigating the association between reporting results completely for benefit outcomes and the statistical significance of the results. Data comes from Chan et al.<sup>23,75</sup> and Dwan et al.<sup>29</sup>.

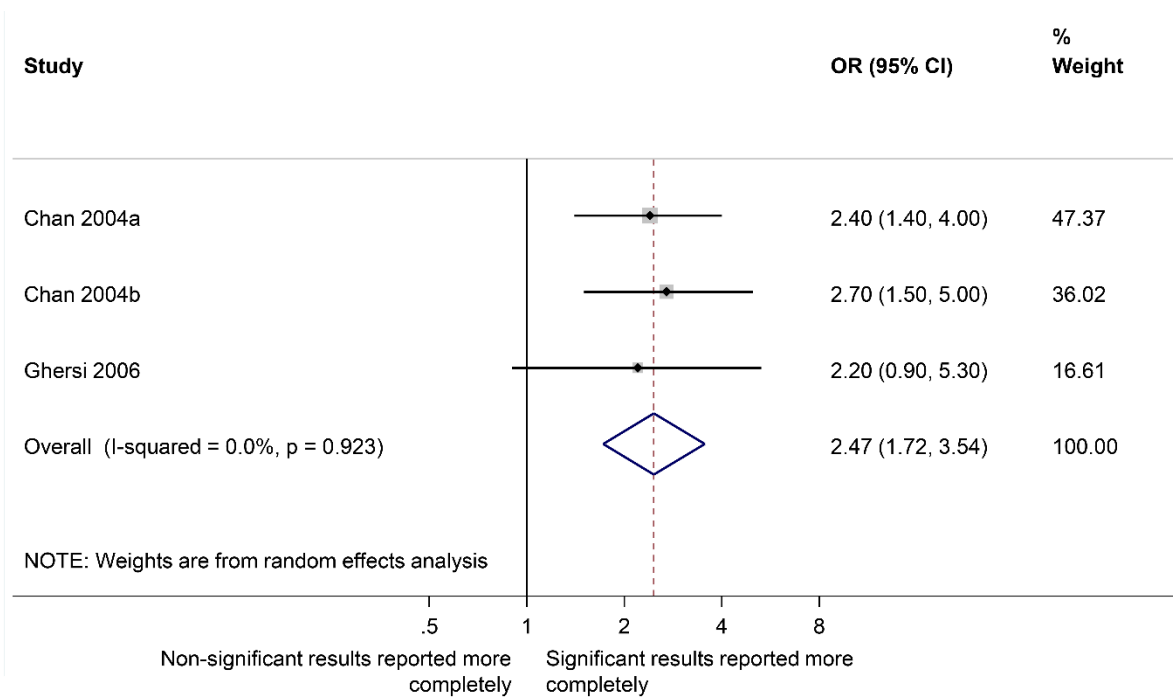




Figure 3. Hypothetical funnel plots: (A) symmetrical plot in the absence of reporting bias (open circles indicate smaller studies showing no statistically significant results); (B) asymmetrical plot in the presence of reporting bias (smaller studies showing no statistically significant results are missing); (C) asymmetrical plot in the presence of bias due to methodologically flawed smaller studies (open circles indicate small studies using few methodological safeguards, whose results are biased towards larger effects).

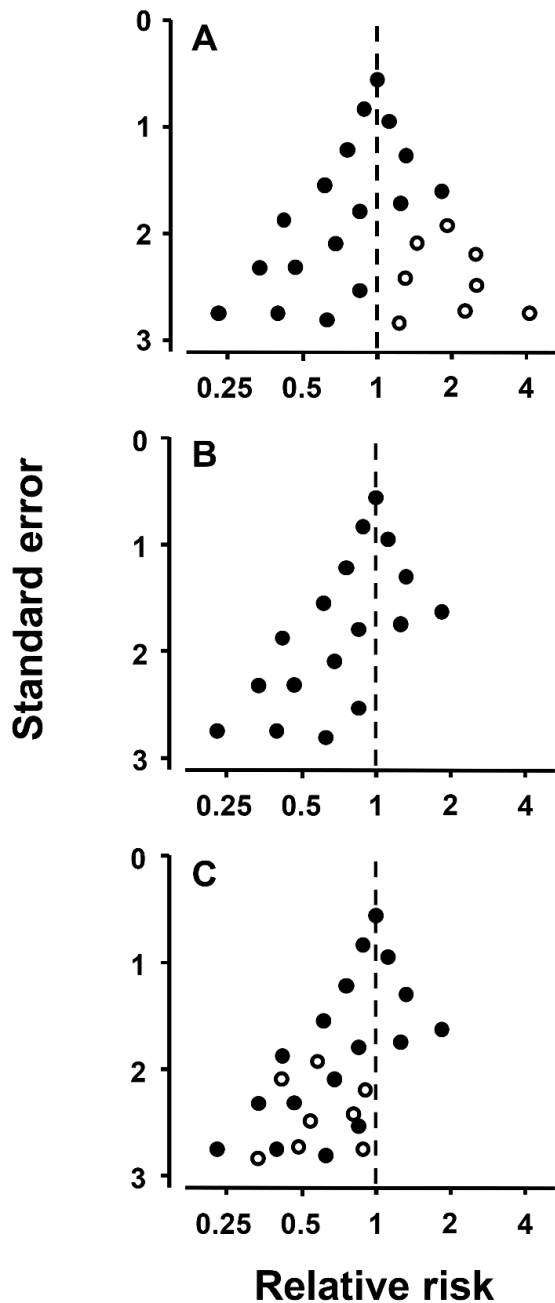


Figure 4. Contour-enhanced funnel plot for meta-analysis of the effect of selective serotonin reuptake inhibitors (SSRIs) versus placebo on treatment response (Clinical Global Impressions Improvement scale [CGI-I])<sup>59</sup>. There is a suggestion of missing results in the left-hand side of the plot, where results would be unfavourable to SSRIs and in the area of statistical nonsignificance (i.e. the white area where  $P > 0.10$ ), which adds credence to the possibility that the asymmetry is due to reporting biases.

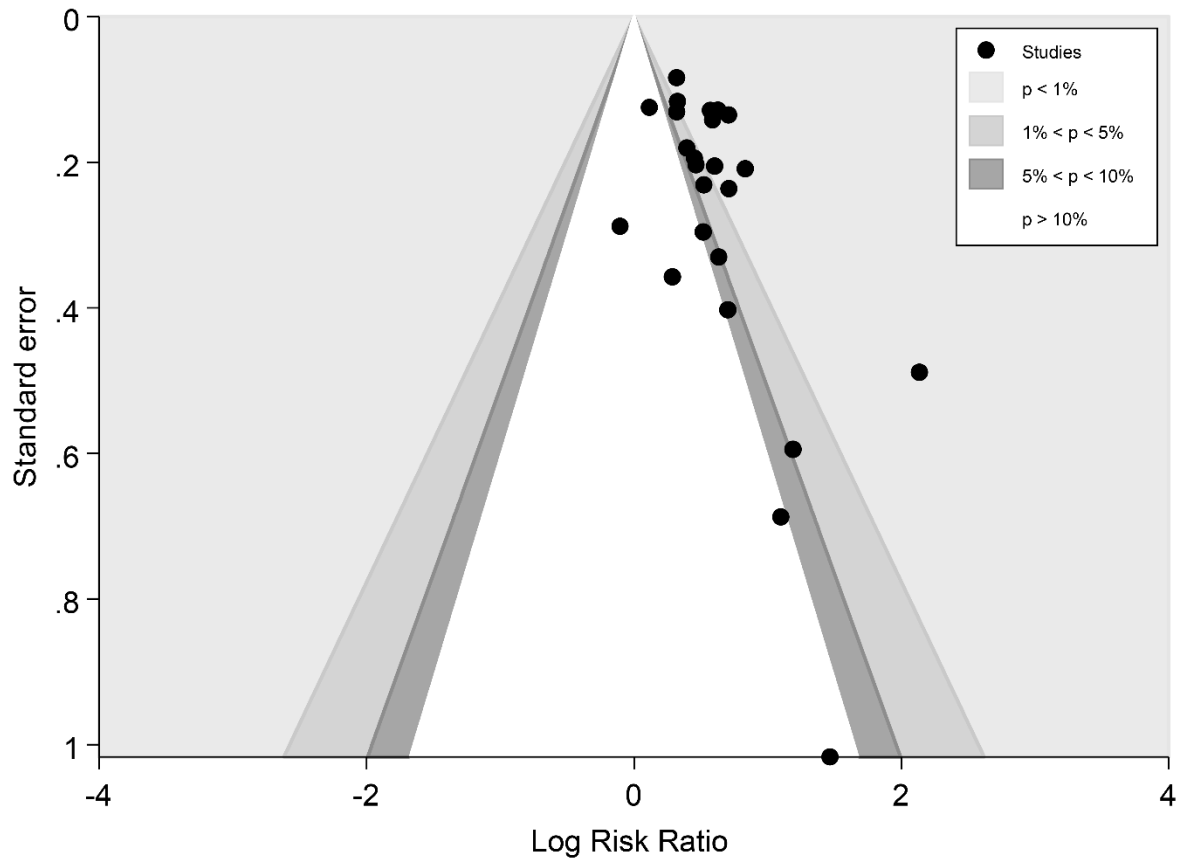


Figure 5. Contour-enhanced funnel plot for meta-analysis of the effect of higher versus lower intake of long-chain omega-3 fats on all-cause mortality<sup>60</sup>. There is a suggestion of missing results in the right-hand side of the plot, where results would be favourable to lower intake of omega-3 and in the area of statistical nonsignificance (i.e. the white area where  $P > 0.10$ ). However, given that almost all results in the plot appear in the area of statistical nonsignificance, this reduces the plausibility that reporting bias is the underlying cause of this funnel plot asymmetry.

