RESEARCH METHODS & REPORTING

Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials

Funnel plots, and tests for funnel plot asymmetry, have been widely used to examine bias in the results of meta-analyses. Funnel plot asymmetry should not be equated with publication bias, because it has a number of other possible causes. This article describes how to interpret funnel plot asymmetry, recommends appropriate tests, and explains the implications for choice of meta-analysis model

Jonathan A C Sterne *professor*[1], Alex J Sutton *professor*[2], John P A Ioannidis *professor and director*[3], Norma Terrin *associate professor*[4], David R Jones *professor*[2], Joseph Lau *professor*[4], James Carpenter *reader*[5], Gerta Rücker *research assistant*[6], Roger M Harbord *research associate*[1], Christopher H Schmid *professor*[4], Jennifer Tetzlaff *research coordinator*[7], Jonathan J Deeks *professor*[8], Jaime Peters *research fellow*[9], Petra Macaskill *associate professor*[10], Guido Schwarzer *research assistant*[6], Sue Duval *assistant professor*[11], Douglas G Altman *professor*[12], David Moher *senior scientist*[7], Julian P T Higgins *senior statistician*[13]

[1]School of Social and Community Medicine, University of Bristol, Bristol BS8 2PS, UK; [2]Department of Health Sciences, University of Leicester, Leicester, UK; [3]Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, CA, USA; [4]Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA, USA; [5]Medical Statistics Unit, London School of Hygiene and Tropical Medicine, London, UK ; [6]Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany; [7]Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada; [8]School of Health and Population Sciences, University of Birmingham, Birmingham, UK; [9]Peninsula Medical School, University of Exeter, Exeter, UK; [10]School of Public Health, University of Sydney, NSW, Australia; [11]University of Minnesota School of Public Health, Minneapolis, MN, USA; [12]Centre for Statistics in Medicine, University of Oxford, Oxford, UK; [13]MRC Biostatistics Unit, Cambridge, UK

The 1997 paper describing the test for funnel plot asymmetry proposed by Egger et al [1] is one of the most cited articles in the history of *BMJ*.[1] Despite the recommendations contained in this and subsequent papers,[2][3] funnel plot asymmetry is often, wrongly, equated with publication or other reporting biases. The use and appropriate interpretation of funnel plots and tests for funnel plot asymmetry have been controversial because of questions about statistical validity,[4] disputes over appropriate interpretation,[3][5][6] and low power of the tests.[2]

This article recommends how to examine and interpret funnel plot asymmetry (also known as small study effects[2]) in meta-analyses of randomised controlled trials. The recommendations are based on a detailed MEDLINE review of literature published up to 2007 and discussions among methodologists, who extended and adapted guidance previously summarised in the *Cochrane Handbook for Systematic Reviews of Interventions.*[7]

What is a funnel plot?

A funnel plot is a scatter plot of the effect estimates from individual studies against some measure of each study's size or precision. The standard error of the effect estimate is often chosen as the measure of study size and plotted on the vertical axis[8] with a reversed scale that places the larger, most powerful studies towards the top. The effect estimates from smaller studies should scatter more widely at the bottom, with the spread narrowing among larger studies.[9] In the absence of bias and between study heterogeneity, the scatter will be due to sampling variation alone and the plot will resemble a symmetrical inverted funnel (fig 1). A triangle centred on a fixed effect summary estimate and extending 1.96 standard errors either side will

Correspondence to: J A C Sterne jonathan.sterne@bristol.ac.uk

include about 95% of studies if no bias is present and the fixed effect assumption (that the true treatment effect is the same in each study) is valid. The appendix on bmj.com discusses choice of axis in funnel plots.

Implications of heterogeneity, reporting bias, and chance

Heterogeneity, reporting bias, and chance may all lead to asymmetry or other shapes in funnel plots (box). Funnel plot asymmetry may also be an artefact of the choice of statistics being plotted (see appendix). The presence of any shape in a funnel plot is contingent on the studies having a range of standard errors, since otherwise they would lie on a horizontal line.

Heterogeneity

Statistical heterogeneity refers to differences between study results beyond those attributable to chance. It may arise because of clinical differences between studies (for example, setting, types of participants, or implementation of the intervention) or methodological differences (such as extent of control over bias). A random effects model is often used to incorporate heterogeneity in meta-analyses. If the heterogeneity fits with the assumptions of this model, a funnel plot will be symmetrical but with additional horizontal scatter. If heterogeneity is large it may overwhelm the sampling error, so that the plot appears cylindrical.

Heterogeneity will lead to funnel plot asymmetry if it induces a correlation between study sizes and intervention effects.[5] For example, substantial benefit may be seen only in high risk patients, and these may be preferentially included in early, small studies.[10] Or the intervention may have been implemented less thoroughly in larger studies, resulting in smaller effect estimates compared with smaller studies.[11]

Figure 2 shows funnel plot asymmetry arising from heterogeneity that is due entirely to there being three distinct subgroups of studies, each with a different intervention effect.[12] The separate funnels for each subgroup are symmetrical. Unfortunately, in practice, important sources of heterogeneity are often unknown.

Differences in methodological quality may also cause heterogeneity and lead to funnel plot asymmetry. Smaller studies tend to be conducted and analysed with less methodological rigour than larger studies,[13] and trials of lower quality also tend to show larger intervention effects.[14] [15]

Reporting bias

Reporting biases arise when the dissemination of research findings is influenced by the nature and direction of results. Statistically significant "positive" results are more likely to be published, published rapidly, published in English, published more than once, published in high impact journals, and cited by others.[16-19] Data that would lead to negative results may be filtered, manipulated, or presented in such a way that they become positive.[14] [20]

Reporting biases can have three types of consequence for a meta-analysis:

- A systematic review may fail to locate an eligible study because all information about it is suppressed or hard to find (publication bias)
- A located study may not provide usable data for the outcome of interest because the study authors did not

consider the result sufficiently interesting (selective outcome reporting)
- A located study may provide biased results for some outcome—for example, by presenting the result with the smallest P value or largest effect estimate after trying several analysis methods (selective analysis reporting).

These biases may cause funnel plot asymmetry if statistically significant results suggesting a beneficial effect are more likely to be published than non-significant results. Such asymmetry may be exaggerated if there is a further tendency for smaller studies to be more prone to selective suppression of results than larger studies. This is often assumed to be the case for randomised trials. For instance, it is probably more difficult to make a large study disappear without trace, while a small study can easily be lost in a file drawer.[21] The same may apply to specific outcomes—for example, it is difficult not to report on mortality or myocardial infarction if these are outcomes of a large study.

Smaller studies have more sampling error in their effect estimates. Thus even though the risk of a false positive significant finding is the same, multiple analyses are more likely to yield a large effect estimate that may seem worth publishing. However, biases may not act this way in real life; funnel plots could be symmetrical even in the presence of publication bias or selective outcome reporting[19] [22]—for example, if the published findings point to effects in different directions but unreported results indicate neither direction. Alternatively, bias may have affected few studies and therefore not cause glaring asymmetry.

Chance

The role of chance is critical for interpretation of funnel plots because most meta-analyses of randomised trials in healthcare contain few studies.[2] Investigations of relations across studies in a meta-analysis are seriously prone to false positive findings when there is a small number of studies and heterogeneity across studies,[23] and this may affect funnel plot symmetry.

Interpreting funnel plot asymmetry

Authors of systematic reviews should distinguish between possible reasons for funnel plot asymmetry (box 1). Knowledge of the intervention, and the circumstances in which it was implemented in different studies, can help identify causes of asymmetry in funnel plots, which should also be interpreted in the context of susceptibility to biases of research in the field of interest. Potential conflicts of interest, whether outcomes and analyses have been standardised, and extent of trial registration may need to be considered. For example, studies of antidepressants generate substantial conflicts of interest because the drugs generate vast sales revenues. Furthermore, there are hundreds of outcome scales, analyses can be very flexible, and trial registration was uncommon until recently.[24] Conversely, in a prospective meta-analysis where all data are included and all analyses fully standardised and conducted according to a predetermined protocol, publication or reporting biases cannot exist. Reporting bias is therefore more likely to be a cause of an asymmetric plot in the first situation than in the second.

Terrin et al found that researchers were poor at identifying publication bias from funnel plots.[5] Including contour lines corresponding to perceived milestones of statistical significance (P=0.01, 0.05, 0.1, etc) may aid visual interpretation.[25] If studies seem to be missing in areas of non-significance (fig 3, top) then asymmetry may be due to reporting bias, although other explanations should still be considered. If the supposed missing

Box 1: Possible sources of asymmetry in funnel plots (adapted from Egger et al[1])

Reporting biases
- Publication bias:
  Delayed publication (also known as time lag or pipeline) bias

  Location biases (eg, language bias, citation bias, multiple publication bias)
- Selective outcome reporting
- Selective analysis reporting

Poor methodological quality leading to spuriously inflated effects in smaller studies
- Poor methodological design
- Inadequate analysis
- Fraud

True heterogeneity
- Size of effect differs according to study size (eg, because of differences in the intensity of interventions or in underlying risk between studies of different sizes)

Artefactual
- In some circumstances, sampling variation can lead to an association between the intervention effect and its standard error

Chance
- Asymmetry may occur by chance, which motivates the use of asymmetry tests

studies are in areas of higher significance or in a direction likely to be considered desirable to their authors (fig 3, bottom), asymmetry is probably due to factors other than reporting bias.

Statistical tests for funnel plot asymmetry

A test for funnel plot asymmetry (sometimes referred to as a test for small study effects) examines whether the association between estimated intervention effects and a measure of study size is greater than might be expected to occur by chance. These tests typically have low power, so even when a test does not provide evidence of asymmetry, bias cannot be excluded. For outcomes measured on a continuous scale a test based on a weighted linear regression of the effect estimates on their standard errors is straightforward.[1] When outcomes are dichotomous and intervention effects are expressed as odds ratios, this corresponds to an inverse variance weighted linear regression of the log odds ratio on its standard error.[2] Unfortunately, there are statistical problems because the standard error of the log odds ratio is mathematically linked to the size of the odds ratio, even in the absence of small study effects.[2 4] Many authors have therefore proposed alternative tests (see appendix on bmj.com).[4 26-28]

Because it is impossible to know the precise mechanism(s) leading to funnel plot asymmetry, simulation studies (in which tests are evaluated on large numbers of computer generated datasets) are required to evaluate test characteristics. Most have examined a range of assumptions about the extent of reporting bias by selectively removing studies from simulated datasets.[26-28] After reviewing the results of these studies, and based on theoretical considerations, we formulated recommendations on testing for funnel plot asymmetry (box 2). The appendix describes the proposed tests, explains the reasons that some were not recommended, and discusses funnel plots for intervention effects measured as risk ratios, risk differences, and standardised mean differences. Our recommendations imply that tests for funnel plot asymmetry should be used in only a minority of meta-analyses.[29]

Funnel plots and meta-analysis models
Fixed and random effects models

Funnel plots can help guide choice of meta-analysis method. Random effects meta-analyses weight studies relatively more equally than fixed effect analyses by incorporating the between study variance into the denominator of each weight. If effect estimates are related to standard errors (funnel plot asymmetry), the random effects estimate will be pulled more towards findings from smaller studies than the fixed effect estimate will be. Random effects models can thus have undesirable consequences and are not always conservative.[30]

The trials of intravenous magnesium after myocardial infarction provide an extreme example of the differences between fixed and random effects analyses that can arise in the presence of funnel plot asymmetry.[31] Beneficial effects on mortality, found in a meta-analysis of small studies,[32] were subsequently contradicted when the very large ISIS-4 study found no evidence of benefit.[33] A contour enhanced funnel plot (fig 4) gives a clear visual impression of asymmetry, which is confirmed by small P values from the Harbord and Peters tests (P<0.001 and P=0.002 respectively).

Figure 5 shows that in a fixed effect analysis ISIS-4 receives 90% of the weight, and there is no evidence of a beneficial effect. However, there is clear evidence of between study heterogeneity (P<0.001, I²=68%), and in a random effects analysis the small studies dominate so that intervention appears beneficial. To interpret the accumulated evidence, it is necessary to make a judgment about the validity or relevance of the combined evidence from the smaller studies compared with that from ISIS-4. The contour enhanced funnel plot suggests that publication bias does not completely explain the asymmetry, since many of the beneficial effects reported from smaller studies were not significant. Plausible explanations for these results are that methodological flaws in the smaller studies, or changes in the standard of care (widespread adoption of treatments such as aspirin, heparin, and thrombolysis), led to apparent beneficial effects of magnesium. This belief was

Box 2: Recommendations on testing for funnel plot asymmetry

All types of outcome

- As a rule of thumb, tests for funnel plot asymmetry should not be used when there are fewer than 10 studies in the meta-analysis because test power is usually too low to distinguish chance from real asymmetry. (The lower the power of a test, the higher the proportion of "statistically significant" results in which there is in reality no association between study size and intervention effects). In some situations—for example, when there is substantial heterogeneity—the minimum number of studies may be substantially more than 10

- Test results should be interpreted in the context of visual inspection of funnel plots— for example, are there studies with markedly different intervention effect estimates or studies that are highly influential in the asymmetry test? Even if an asymmetry test is statistically significant, publication bias can probably be excluded if small studies tend to lead to lower estimates of benefit than larger studies or if there are no studies with significant results

- When there is evidence of funnel plot asymmetry, publication bias is only one possible explanation (see box 1)

- As far as possible, testing strategy should be specified in advance: choice of test may depend on the degree of heterogeneity observed. Applying and reporting many tests is discouraged: if more than one test is used, all test results should be reported

- Tests for funnel plot asymmetry should not be used if the standard errors of the intervention effect estimates are all similar (the studies are of similar sizes)

Continuous outcomes with intervention effects measured as mean differences

- The test proposed by Egger et al may be used to test for funnel plot asymmetry.[1] There is no reason to prefer more recently proposed tests, although their relative advantages and disadvantages have not been formally examined. General considerations suggest that the power will be greater than for dichotomous outcomes but that use of the test with substantially fewer than 10 studies would be unwise

Dichotomous outcomes with intervention effects measured as odds ratios

- The tests proposed by Harbord et al[26] and Peters et al[27] avoid the mathematical association between the log odds ratio and its standard error when there is a substantial intervention effect while retaining power compared with alternative tests. However, false positive results may still occur if there is substantial between study heterogeneity

- If there is substantial between study heterogeneity (the estimated heterogeneity variance of log odds ratios, τ^2, is >0.1) only the arcsine test including random effects, proposed by Rücker et al, has been shown to work reasonably well.[28] However, it is slightly conservative in the absence of heterogeneity and its interpretation is less familiar than for other tests because it is based on an arcsine transformation.

- When τ^2 is <0.1, one of the tests proposed by Harbord et al,[26] Peters et al,[27] or Rücker et al[28] can be used. Test performance generally deteriorates as τ^2 increases.

reinforced by the subsequent publication of the MAGIC trial, in which magnesium added to these treatments which also found no evidence of benefit on mortality (odds ratio 1.0, 95% confidence interval 0.8 to 1.1).[34]

We recommend that when review authors are concerned about funnel plot asymmetry in a meta-analysis with evidence of between study heterogeneity, they should compare the fixed and random effects estimates of the intervention effect. If the random effects estimate is more beneficial, authors should consider whether it is plausible that the intervention is more effective in smaller studies. Formal investigations of heterogeneity of effects may reveal explanations for funnel plot asymmetry, in which case presentation of results should focus on these. If larger studies tend to be methodologically superior to smaller studies, or were conducted in circumstances more typical of the use of the intervention in practice, it may be appropriate to include only larger studies in the meta-analysis.

Extrapolation of a funnel plot regression line

An assumed relation between susceptibility to bias and study size can be exploited by extrapolating within a funnel plot. When funnel plot asymmetry is due to bias rather than substantive heterogeneity, it is usually assumed that results from larger studies are more believable than those from smaller studies because they are less susceptible to methodological flaws or reporting biases. Extrapolating a regression line on a funnel plot to minimum bias (maximum sample size) produces a meta-analytical estimate that can be regarded as corrected for such biases.[35 36 37] However, because it is difficult to distinguish

between asymmetry due to bias and asymmetry due to heterogeneity or chance, the broad applicability of such approaches is uncertain. Further approaches to adjusting for publication bias are described and discussed in the appendix.

Discussion

Reporting biases are one of a number of possible explanations for the associations between study size and effect size that are displayed in asymmetric funnel plots. Examining and testing for funnel plot asymmetry, when appropriate, is an important means of addressing bias in meta-analyses, but the multiple causes of asymmetry and limited power of asymmetry tests mean that other ways to address reporting biases are also of importance. Searches of online trial registries can identify unpublished trials, although they do not currently guarantee access to trial protocols and results. When there are no registered but unpublished trials, and the outcome of interest is reported by all trials, restricting meta-analyses to registered trials should preclude publication bias. Recent comparisons of results of published trials with those submitted for regulatory approval have also provided clear evidence of reporting bias.[38 39] Methods for dealing with selective reporting of outcomes have been described elsewhere.[40]

Our recommendations apply to meta-analyses of randomised trials, and their applicability in other contexts such as meta-analyses of epidemiological or diagnostic test studies is unclear.[41] The performance of tests for funnel plot asymmetry in these contexts is likely to differ from that in meta-analyses of randomised trials. Further factors, such as confounding and

precision of measurements, may cause a relation between study size and effect estimates in observational studies. For example, large studies based on routinely collected data might not fully control confounding compared with smaller, purpose designed studies that collected a wide range of potential confounding variables. Alternatively, larger studies might use self reported exposure levels, which are more error prone, while smaller studies used precise measuring instruments. However, simulation studies have usually not considered such situations. An exception is for diagnostic studies, where large imbalances in group sizes and substantial odds ratios lead to poor performance of some tests: that proposed by Deeks et al was designed for use in this context.[4]

1. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629-34.
2. Sterne JAC, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J Clin Epidemiol* 2000;53:1119-29.
3. Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. *BMJ* 2006;333:597-600.
4. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005;58:882-93.
5. Terrin N, Schmid CH, Lau J. In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *J Clin Epidemiol* 2005;58:894-901.
6. Ioannidis JP. Interpretation of tests of heterogeneity and bias in meta-analysis. *J Eval Clin Pract* 2008;14:951-7.
7. Sterne JAC, Egger M, Moher D. Addressing reporting biases. In: Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions* . Wiley, 2008.
8. Sterne JAC, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J Clin Epidemiol* 2001;54:1046-55.
9. Begg CB, Berlin JA. Publication bias: a problem in interpreting medical data. *J R Statist Soc A* 1988;151:419-63.
10. Davey Smith G, Egger M. Who benefits from medical interventions? Treating low risk patients can be a high risk strategy. *BMJ* 1994;308:72-4.
11. Stuck AE, Siu AL, Wieland GD, Adams J, Rubenstein LZ. Comprehensive geriatric assessment: a meta-analysis of controlled trials. *Lancet* 1993;342:1032-6.
12. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L, Moreno SG. Assessing publication bias in meta-analyses in the presence of between-study heterogeneity. *J R Statist Soc A* 2010;173:575-91.
13. Egger M, Jüni 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 Technol Assess* 2003;7:1-68.
14. Ioannidis JP. Why most discovered true associations are inflated. *Epidemiology* 2008;19:640-8.
15. Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;336:601-5.
16. Hopewell S, Clarke M, Stewart L, Tierney J. Time to publication for results of clinical trials. *Cochrane Database Syst Rev* 2007;2:MR000011.
17. Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database Syst Rev* 2009;1:MR000006.
18. Song F, Parekh S, Hooper L, Loke YK, Ryder J, Sutton J, et al. Dissemination and publication of research findings: an updated review of related biases. *Health Technol Assess* 2010;14:iii,ix-iii,193.
19. Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan AW, Cronin E, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS ONE* 2008;3:e3081.
20. 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.
21. Rosenthal R. The "file drawer" problem and tolerance for null results. *Psychol Bull* 1979;86:638-41.
22. 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;291:2457-65.
23. Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. *Stat Med* 2004;23:1663-82.
24. Ioannidis JP. Effectiveness of antidepressants: an evidence myth constructed from a thousand randomized trials? *Philos Ethics Humanit Med* 2008;3:14.
25. Peters J, Sutton AJ, Jones DR, Abrams KR, Rushton L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J Clin Epidemiol* 2008;61:991-6.
26. 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.
27. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of two methods to detect publication bias in meta-analysis. *JAMA* 2006;295:676-80.
28. Rücker G, Schwarzer G, Carpenter J. Arcsine test for publication bias in meta-analyses with binary outcomes. *Stat Med* 2008;27:746-63.
29. Ioannidis JP, Trikalinos TA. The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. *CMAJ* 2007;176:1091-6.
30. Poole C, Greenland S. Random-effects meta-analyses are not always conservative. *Am J Epidemiol* 1999;150:469-75.
31. Egger M, Davey Smith G. Misleading meta-analysis. Lessons from an "effective, safe, simple" intervention that wasn't. *BMJ* 1995;310:752-4.
32. Teo KK, Yusuf S, Collins R, Held PH, Peto R. Effects of intravenous magnesium in suspected acute myocardial infarction: overview of randomised trials. *BMJ* 1991;303:1499-503.
33. ISIS-4 (Fourth International Study of Infarct Survival) Collaborative Group. ISIS-4: a randomised factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58,050 patients with suspected acute myocardial infarction. *Lancet* 1995;345:669-85.
34. Early administration of intravenous magnesium to high-risk patients with acute myocardial infarction in the Magnesium in Coronaries (MAGIC) Trial: a randomised controlled trial. *Lancet* 2002;360:1189-96.
35. Shang A, Huwiler-Muntener K, Nartey L, Jüni P, Dörig S, Stene JA, et al. Are the clinical effects of homoeopathy placebo effects? Comparative study of placebo-controlled trials of homoeopathy and allopathy. *Lancet* 2005;366:726-32.
36. Moreno SG, Sutton AJ, Ades AE, Stanley TD, Abrams KR, Peters JL, et al. Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BMC Med Res Methodol* 2009;9:2.
37. Rucker G, Schwarzer G, Carpenter JR, Binder H, Schumacher M. Treatment-effect estimates adjusted for small-study effects via a limit meta-analysis. *Biostatistics* 2011;12:122-42.
38. Moreno SG, Sutton AJ, Turner EH, Abrams KR, Cooper NJ, Palmer TM, 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.
39. Eyding D, Lelgemann M, Grouven U, Härter M, Kromp M, Kaiser T, et al. Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials. *BMJ* 2010;341:c4737.
40. Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 2010;340:c365.
41. Egger M, Schneider M, Davey Smith G. Spurious precision? Meta-analysis of observational studies. *BMJ* 1998;316:140-4.

Cite this as: *BMJ* 2011;342:d4002

Summary points

Inferences on the presence of bias or heterogeneity should consider different causes of funnel plot asymmetry and should not be based on visual inspection of funnel plots alone

They should be informed by contextual factors, including the plausibility of publication bias as an explanation for the asymmetry

Testing for funnel plot asymmetry should follow the recommendations detailed in this article

The fixed and random effects estimates of the intervention effect should be compared when funnel plot asymmetry exists in a meta-analysis with between study heterogeneity

Figures
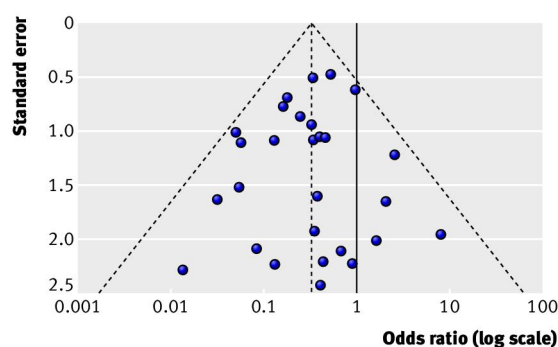
Fig 1 Example of symmetrical funnel plot. The outer dashed lines indicate the triangular region within which 95% of studies are expected to lie in the absence of both biases and heterogeneity (fixed effect summary log odds ratio±1.96×standard error of summary log odds ratio). The solid vertical line corresponds to no intervention effect
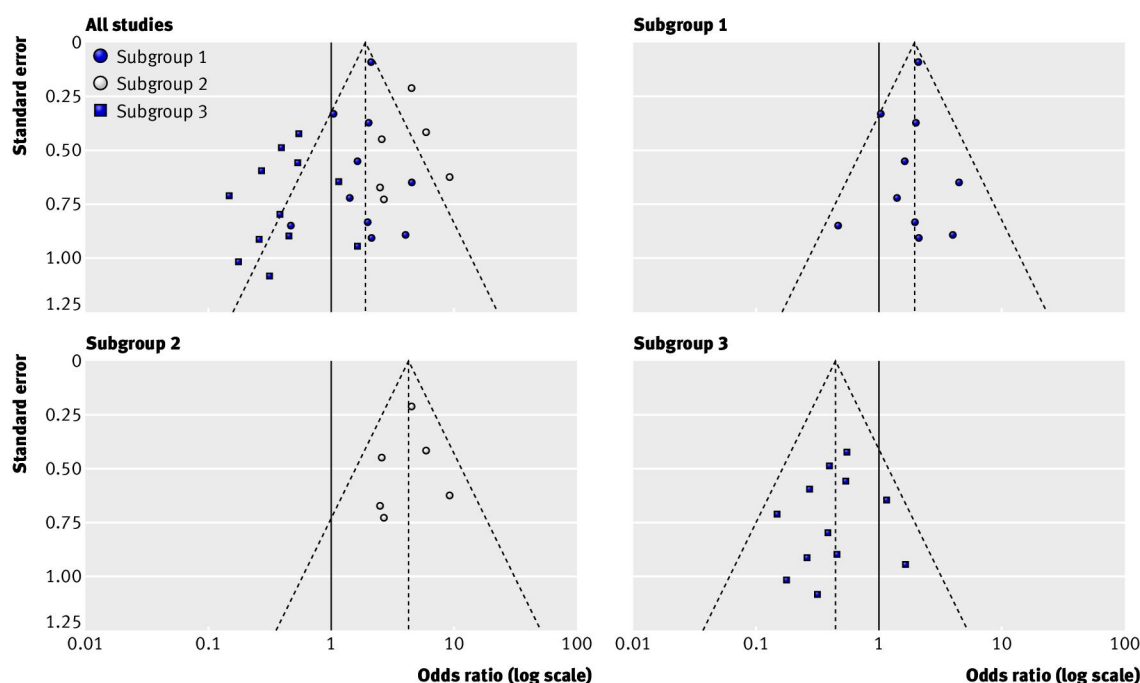

Fig 2 Illustration of funnel plot asymmetry due to heterogeneity, in the form of three distinct subgroups of studies. Funnel plot including all studies (top left) shows clear asymmetry (P<0.001 from Egger test for funnel plot asymmetry). P values for each subgroup are all >0.49.
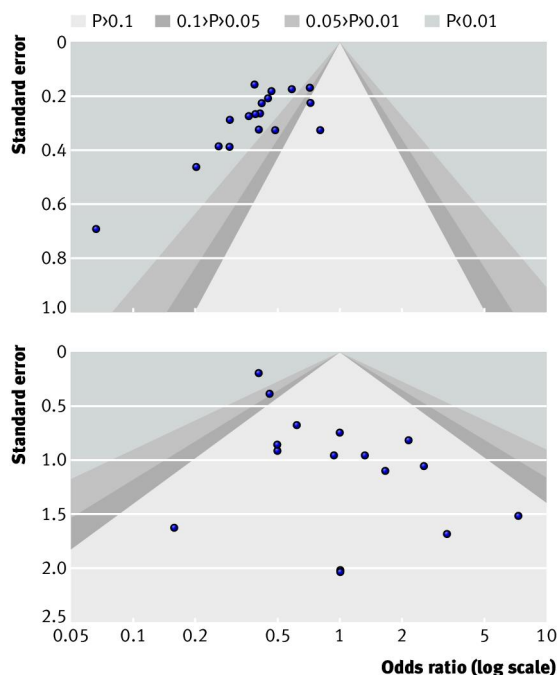
Fig 3 Contour enhanced funnel plots. In the top diagram there is a suggestion of missing studies in the middle and right of the plot, broadly in the white area of non-significance, making publication bias plausible. In the bottom diagram there is a suggestion of missing studies on the bottom left hand side of the plot. Since most of this area contains regions of high significance, publication bias is unlikely to be the underlying cause of asymmetry
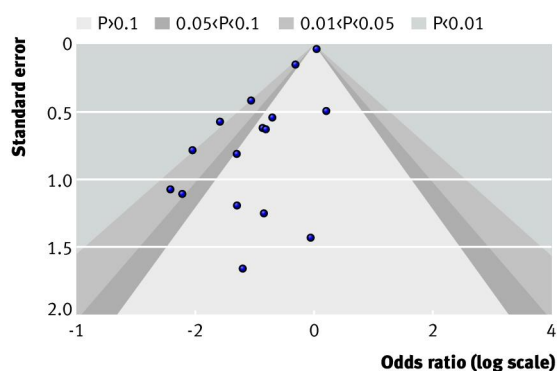

Fig 4 Contour enhanced funnel plot for trials of the effect of intravenous magnesium on mortality after myocardial infarction
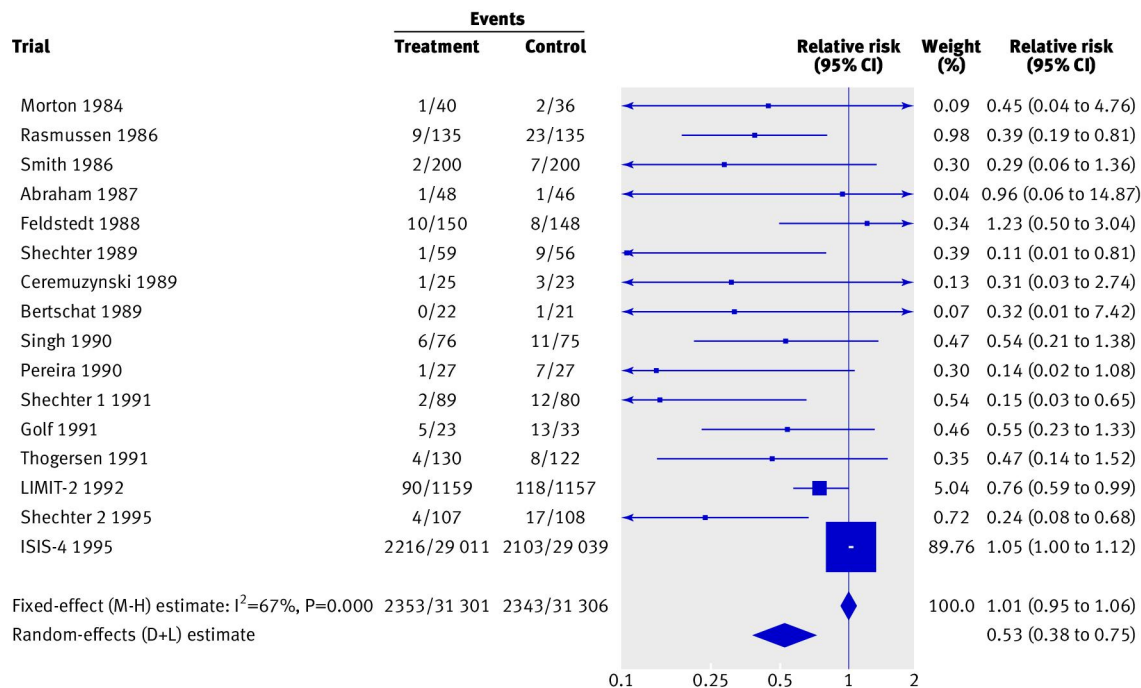
Fig 5 Comparison of fixed and random effects meta-analytical estimates of the effect of intravenous magnesium on mortality after myocardial infarction