Bias in meta-analysis detected by a simple, graphical test

Matthias Egger, George Davey Smith, Martin Schneider, Christoph Minder

Abstract

Objective: Funnel plots (plots of effect estimates against sample size) may be useful to detect bias in meta-analyses that were later contradicted by large trials. We examined whether a simple test of asymmetry of funnel plots predicts discordance of results when meta-analyses are compared to large trials, and we assessed the prevalence of bias in published meta-analyses.

Design: Medline search to identify pairs consisting of a meta-analysis and a single large trial (concordance of results was assumed if effects were in the same direction and the meta-analytic estimate was within 30% of the trial); analysis of funnel plots from 37 meta-analyses identified from a hand search of four leading general medicine journals 1993-6 and 38 meta-analyses from the second 1996 issue of the *Cochrane Database of Systematic Reviews*.

Main outcome measure: Degree of funnel plot asymmetry as measured by the intercept from regression of standard normal deviates against precision.

Results: In the eight pairs of meta-analysis and large trial that were identified (five from cardiovascular medicine, one from diabetic medicine, one from geriatric medicine, one from perinatal medicine) there were four concordant and four discordant pairs. In all cases discordance was due to meta-analyses showing larger effects. Funnel plot asymmetry was present in three out of four discordant pairs but in none of concordant pairs. In 14 (38%) journal meta-analyses and 5 (13%) Cochrane reviews, funnel plot asymmetry indicated that there was bias.

Conclusions: A simple analysis of funnel plots provides a useful test for the likely presence of bias in meta-analyses, but as the capacity to detect bias will be limited when meta-analyses are based on a limited number of small trials the results from such analyses should be treated with considerable caution.

Introduction

Systematic reviews of the best available evidence regarding the benefits and risks of medical interventions can inform decision making in clinical practice and public health.[1,2] Such reviews are, whenever possible, based on meta-analysis: "a statistical analysis which combines or integrates the results of several independent clinical trials considered by the analyst to be 'combinable.' "[3] However, the findings of some meta-analyses have later been contradicted by large randomised controlled trials.[4] Such discrepancies have brought discredit on a technique that has been controversial since the outset.[5] The appearance of misleading meta-analysis is not surprising considering the existence of publication bias and the many other biases that may be introduced in the process of locating, selecting, and combining studies.[6-9]

Funnel plots, plots of the trials' effect estimates against sample size, may be useful to assess the validity of meta-analyses.[4,10] The funnel plot is based on the fact that precision in estimating the underlying treatment effect will increase as the sample size of component studies increases. Results from small studies will scatter widely at the bottom of the graph, with the spread narrowing among larger studies. In the absence of bias the plot will resemble a symmetrical inverted funnel. Conversely, if there is bias, funnel plots will often be skewed and asymmetrical.

The value of the funnel plot has not been systematically examined, and symmetry (or asymmetry) has generally been defined informally, through visual examination. Unsurprisingly, funnel plots have been interpreted differently by different observers.[11] We measured funnel plot asymmetry numerically and examined the extent to which such asymmetry predicts discordance of results when meta-analyses are compared to single large trials of the same issue. We used the same method to assess the prevalence of funnel plot asymmetry, and thus of possible bias, among meta-analyses published in leading general medicine journals and meta-analyses disseminated electronically by the Cochrane Collaboration.

Methods

Measures of funnel plot asymmetry

We used a linear regression approach to measure funnel plot asymmetry on the natural logarithm scale of the odds ratio. This corresponds to a regression analysis of Galbraith's radial plot,[12] although in the present context the regression is not constrained to run through the origin. The standard normal deviate (SND), defined as the odds ratio divided by its standard error, is regressed against the estimate's precision, the latter being defined as the inverse of the standard error (regression equation: $\text{SND} = a + b \times \text{precision}$). As precision depends largely on sample size, small trials will be close to zero on the x axis. Small trials may produce an odds ratio that differs from unity, but because the

Department of Social Medicine, University of Bristol, Bristol BS8 2PR

Matthias Egger, *reader in social medicine and epidemiology*

George Davey Smith, *professor of clinical epidemiology*

Department of Social and Preventive Medicine, University of Berne, CH-3012 Berne, Switzerland

Martin Schneider, *research associate*

Christoph Minder, *head, medical statistics unit*

Correspondence to: Dr Egger m.egger@bristol.ac.uk

BMJ 1997;315:629–34

standard error will be large, the resulting standard normal deviate will again be close to zero. Small trials will thus be close to zero on both axes—that is, close to the origin. Conversely, large studies will produce precise estimates and, if the treatment is effective, also produce large standard normal deviates. The points from a homogeneous set of trials, not distorted by selection bias, will thus scatter about a line that runs through the origin at standard normal deviate zero ($a=0$), with the slope b indicating the size and direction of effect.[12] This situation corresponds to a symmetrical funnel plot.

If there is asymmetry, with smaller studies showing effects that differ systematically from larger studies, the regression line will not run through the origin. The intercept a provides a measure of asymmetry—the larger its deviation from zero the more pronounced the asymmetry. If the smaller studies show big protective effects, they will force the regression line below the origin on the logarithmic scale. Negative values will therefore indicate that smaller studies show more pronounced beneficial effects than larger studies. In some situations (for example, if there are several small trials but only one larger study) power is gained by weighting the analysis by the inverse of the variance of the effect estimate. We performed both weighted and unweighted analyses and used the output from the analysis yielding the intercept with the larger deviation from zero.

In contrast to the overall test of heterogeneity, the test for funnel plot asymmetry assesses a specific type of heterogeneity and provides a more powerful test in this situation. However, any analysis of heterogeneity depends on the number of trials included in a meta-analysis, which is generally small, and this limits the statistical power of the test. We therefore based evidence of asymmetry on $P<0.1$, and we present intercepts with 90% confidence intervals. The same significance level has been used in previous analyses of heterogeneity in meta-analysis.[13 14]

Identification of meta-analyses and matching large randomised trials

A Medline search (Knight Ridder Information Services, Berne, Switzerland) covering the period January 1985 to April 1996 was performed in April 1996 to identify published meta-analyses. For this purpose the word "meta-analysis" was entered in a free text search. The articles identified included all those indexed with the Medical Subject Heading (MeSH) keyword "meta-analysis," which was introduced in 1989, and articles without the keyword which carried the word meta-analysis in their title or abstract. Results were tabulated by source of publication, and the items published in journals which yielded 30 or more hits were examined further. Meta-analyses of controlled trials combining at least five trials with binary endpoints were identified.

Large scale randomised controlled trials of the same interventions which had been published after the meta-analyses were identified by a Medline search using appropriate keywords. Large trials had to provide an effect estimate with a precision of at least 5. For example, a trial among patients with heart failure in which mortality in the control group at three months is 5%[15] and in which mortality is reduced to 3% among treated patients will need to randomise 2800 patients to measure this effect with a precision of 5 and about 12 000 patients for a precision of 10. Also, the effect estimate from the large trials had to be of equal or greater precision than the meta-analysis. We scrutinised potential matching pairs of meta-analyses and large trials with regard to study participants, interventions, end points and lengths of follow up. In some cases a further Medline search was performed to identify a meta-analysis published in any journal indexed in Medline which would be more suitable for comparison with the large trial.

Some meta-analyses were published several years before the corresponding large trial. In these cases we examined whether the shape of the funnel plot changed when the meta-analysis was updated with trials published in the intervening period.

Concordance and discordance of results

Comparison of results from meta-analyses and large trials required expressing results on a common scale. Odds ratios were used for this purpose. The meta-analysis and the large trial were considered concordant when effects were in the same direction and the estimates from the meta-analysis were within 30% of the estimate of the single trial. A difference of 30% was proposed by Villar et al to denote high similarity between the results from meta-analyses and large trials.[11]

SAS version 6.11 software package (Statistical Analysis System, Cary, NC) was used for statistical analysis.

Table 1 Characteristics of nine pairs of meta-analyses and corresponding large trials

Topic	Typical end point	Large trial (year of publication)	Meta-analysis (year of publication)
Concordant pairs			
β Blockers in myocardial infarction	Mortality in hospital	ISIS-1 (1986)[20]	Yusuf et al (1985)[19]
Streptokinase in myocardial infarction	Mortality in hospital	GISSI-1 (1986)[18]	Yusuf et al (1985)[17]
Angiotensin converting enzyme inhibitors in heart failure	Mortality at 3 months	SOLVD (1991)[15]	Mulrow et al (1988)[26]
Intensive therapy in insulin-dependent diabetes mellitus	Progression of retinopathy over several years	DCCT (1993)[22]	Wang et al (1993)[21]
Discordant pairs			
Magnesium in myocardial infarction	Mortality in hospital	ISIS-4 (1995)[28]	Teo et al (1993)[27]
Nitrates in myocardial infarction	Mortality in hospital	GISSI-3 (1994)[25]	Yusuf et al (1988)[24]
Inpatient geriatric consultation service	Mortality at 6 months	HMO (1995)[23]	Stuck et al (1993)[14]
Aspirin for preventing pre-eclampsia	Development of pre-eclampsia	CLASP (1994)[30]	Imperiale and Stollenwerk (1991)[29]

GISSI=Gruppo Italiano per lo Studio della Streptochinasi nell'Infarto miocardico; ISIS=International Study of Infarct Survival; DCCT=Diabetes Control and Complications Trial; HMO=Health Maintenance Organisation study; SOLVD=Studies of Left Ventricular Dysfunction; CLASP, Collaborative Low-dose Aspirin Study in Pregnancy.

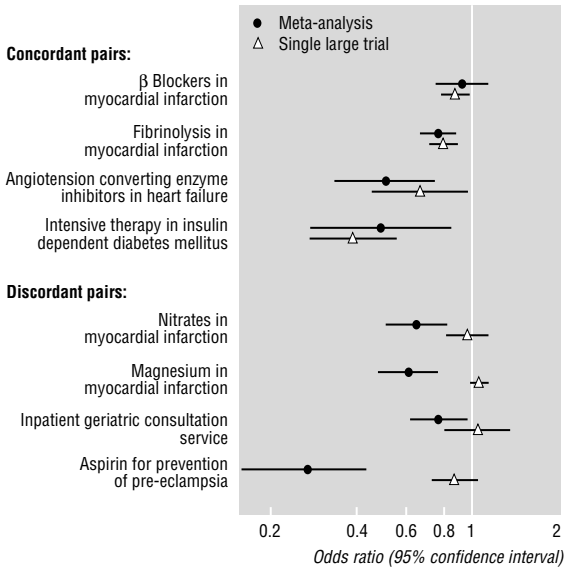

Fig 1 Results from four concordant and four discordant pairs of meta-analysis and large scale randomised controlled trial

Frequency of asymmetry in funnel plots

We performed a hand search of four leading general medicine journals, *Annals of Internal Medicine*, *BMJ*, *JAMA*, and *Lancet*, from 1993 to 1996 and examined the second 1996 issue of the *Cochrane Database of Systematic Reviews*[16] to identify meta-analyses of controlled trials. Analyses that were based on at least five trials with categorical end points were examined further. For each intervention and comparison, the outcome measure which was reported in the largest number of trials was selected. To obtain consistency across reviews, end points were recoded if necessary so that the direction of effect for the expected beneficial outcome was in the same direction. For example, in a review of trials of nicotine patches in smoking cessation, continued smoking rather than quitting was considered to be the outcome, so that an odds ratio above unity indicates an adverse effect.

We identified 38 Cochrane reviews and 37 journal meta-analyses. All references of meta-analyses and trials included are available from the authors on request.

Results

Eight pairs consisting of a meta-analysis and a large trial were identified (table 1).[14 15-30] Five were from cardiovascular medicine, one from diabetic medicine, one from geriatric medicine, and one from perinatal medicine. Effect estimates from meta-analyses had an average precision of 7.9 compared with 14.4 for large trials. There were four concordant pairs[15 17-22 26] and four discordant pairs[14 23-25 27-30] (fig 1). In all cases discordance was a consequence of the meta-analyses showing more beneficial effects than the large trials. Three out of four discordant meta-analyses showed significant ($P < 0.1$) funnel plot asymmetry; funnel plots from concordant pairs showed no significant asymmetry (fig 2, table 2).

Additional trials were identified for three meta-analyses published several years earlier than the large trial.[26 27 29] These were extracted from more recent meta-analyses.[4 31 32] When the meta-analysis of trials of

Table 2 Analysis of funnel plot asymmetry

Meta-analysis	No of trials	Linear regression analysis	
		Intercept (90% CI)	P value
Results concordant with single large trial			
β Blockers in myocardial infarction[19]	26	0.44 (–0.11 to 1.00)	0.19
Streptokinase in myocardial infarction[17]	20	0.59 (–1.30 to 2.48)	0.59
Angiotensin converting enzyme inhibitors in heart failure[26]	13	–0.14 (–0.44 to 0.16)	0.43
Intensive treatment in insulin-dependent diabetes mellitus[21]	6	–0.75 (–2.53 to 1.03)	0.44
Results discordant with single large trial			
Magnesium in myocardial infarction[27]	10	–1.19 (–2.26 to –0.12)	0.068
Nitrates in myocardial infarction[24]	10	–1.84 (–3.25 to –0.43)	0.043
Inpatient geriatric consultation service[14]	8	–2.60 (–4.84 to –0.37)	0.069
Aspirin for preventing pre-eclampsia[29]	6	0.37 (–1.84 to 2.59)	0.75

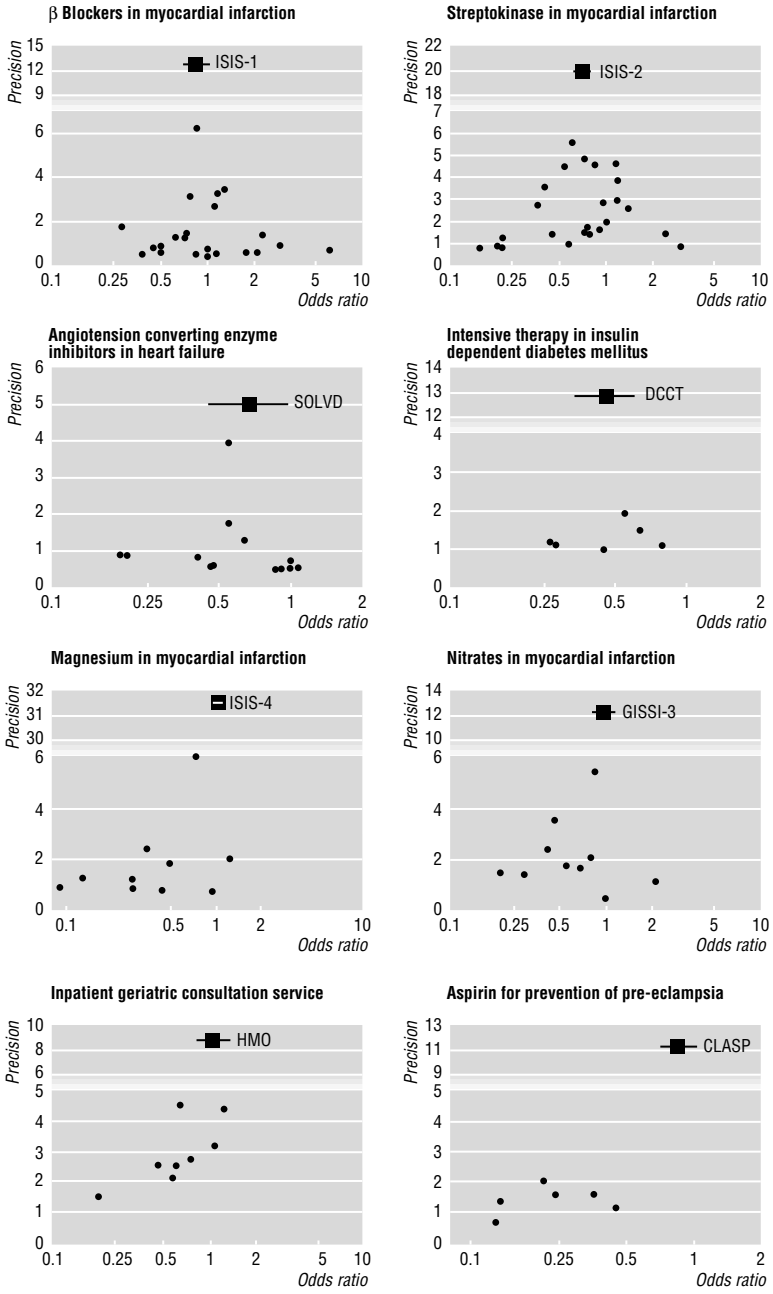

Fig 2 Funnel plots and single large trials. Points indicate odds ratios from trials included in meta-analysis; squares with horizontal lines show odds ratio from large trial with 95% confidence interval. See table 1 for abbreviations of trial names

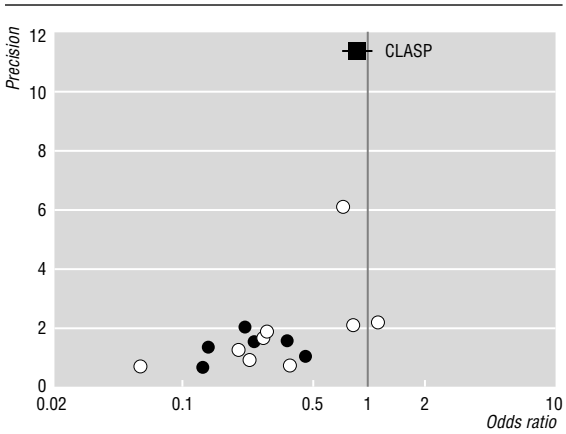

Fig 3 Funnel plot of trials of low dose aspirin in the prevention of pre-eclampsia. Trials included in Imperiale and Stollenwerk's 1991 meta-analysis (closed circles),[29] trials published in subsequent years (1990 to 1993, open circles) and the large 1994 CLASP (collaborative low-dose aspirin study in pregnancy) trial (square with horizontal line indicating 95% confidence interval)[30]

intravenous magnesium in myocardial infarction was updated with five additional trials the intercept indicated even greater asymmetry (−1.36 (90% confidence interval −2.06 to −0.66), P=0.005). When 13 additional trials were added to the analysis of trials of angiotensin converting enzyme inhibitors in heart failure the plot remained symmetrical (intercept 0.07 (−0.53 to 0.67), P=0.85). When the analysis of aspirin for the prevention of pre-eclampsia was updated with nine additional trials, the funnel plot became asymmetrical (intercept −1.49 (−2.20 to −0.79), P=0.003) (fig 3).

Figure 4 shows the distribution of regression intercepts from 38 Cochrane reviews and 37 journal meta-analyses. In the absence of bias, random fluctuations should produce a symmetrical distribution of intercepts around a central value of zero, with an equal number of positive and negative values. This is not what was observed. Distributions were shifted towards negative values, with a mean of −0.24 (−0.65 to 0.17) for Cochrane reviews and −1.00 (−1.50 to −0.49) for journal meta-analyses There were 24 negative and 14 positive intercepts among Cochrane reviews (P=0.10 by sign test) and 26 negative and 11 positive intercepts among journal meta-analyses (P=0.007 by sign test). In five (13%) Cochrane reviews and 14 (38%) journal meta-analyses there was evidence of significant (P<0.1) asymmetry.

Discussion

The selective publication of positive findings from randomised controlled trials is an important concern in meta-analytic reviews of the literature.[9] If the literature is more likely to contain trials showing beneficial effects of treatments, and if equally valid trials showing no effect remain unpublished, how can systematic reviews of this literature serve as an objective guide to decision making in clinical practice and health policy? The potentially serious consequences of such publication bias have been realised for some time, and there have been repeated calls for worldwide registration of clinical trials at inception.[1 4 33-35] Although registration of trials and creation of a database holding the results of both published and unpublished trials would solve the problem, it is unlikely that this will be widely instituted in the foreseeable future.

Critical examination for the presence of publication and related biases must therefore become an essential part of meta-analytic studies and systematic reviews. The findings presented here indicate that a simple graphical and statistical method is useful for this purpose. When testing this method on pairs consisting of meta-analyses and single large trials of the same intervention, we found asymmetry in funnel plots in three out of four pairs with discordant results. The fourth was based on only six trials, and asymmetry emerged when it was updated with further studies.

Sources of funnel plot asymmetry

Publication bias has long been associated with funnel plot asymmetry.[10] Among published studies, however, the probability of identifying relevant trials for meta-analysis is also influenced by their results. English language bias—the preferential publication of "negative" findings in journals published in languages other than English—makes the location and inclusion of such studies less likely.[8] As a consequence of citation bias, "negative" studies are quoted less frequently and are therefore more likely to be missed in the search for relevant trials.[7 36] Results of "positive" trials are sometimes reported more than once, increasing the probability that they will be located for meta-analysis (multiple publication bias).[37] These biases are likely to affect smaller studies to a greater degree than large trials.

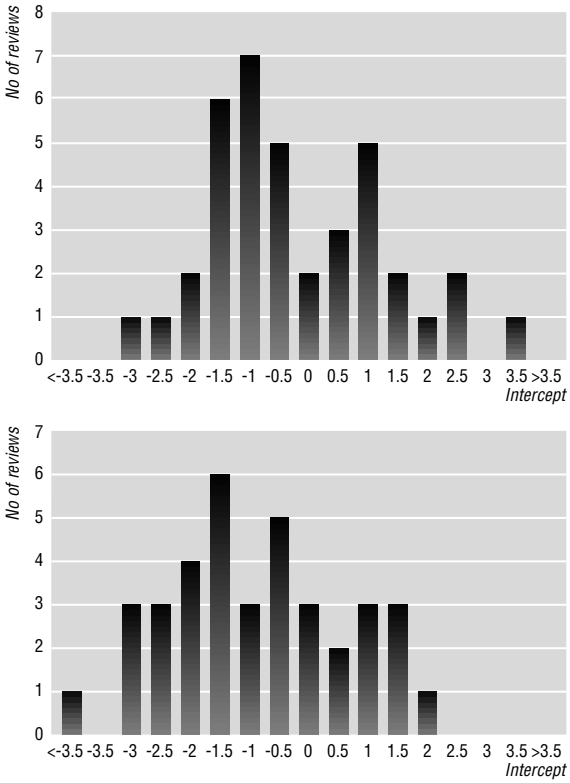

Fig 4 Distribution of intercepts from regression analysis of funnel plot asymmetry for 38 meta-analyses from the *Cochrane Database of Systematic Reviews*, 1996 (upper panel) and 37 meta-analyses published in *Annals of Internal Medicine*, *BMJ*, *JAMA*, and *Lancet* 1993 through 1996 (lower panel)

Sources of asymmetry in funnel plots

Selection bias
- Publication bias
- Location biases:
  - English language bias
  - Citation bias
  - Multiple publication bias

True heterogeneity
- Size of effect differs according to study size:
  - Intensity of intervention
  - Differences in underlying risk

Data irregularities
- Poor methodological design of small studies
- Inadequate analysis
- Fraud

Artefactual
- Choice of effect measure

Chance

Another source of asymmetry arises from differences in methodological quality. Smaller studies are, on average, conducted and analysed with less methodological rigour than larger studies. Trials of lower quality also tend to show the larger effects.[38-40] The degree of symmetry found in a funnel plot may depend on the statistic used to measure effect. Odds ratios overestimate the relative reduction, or increase, in risk if the event rate is high.[41] This can lead to funnel plot asymmetry if the smaller trials were consistently conducted in patients at higher risk. Similarly, if events accrue at a constant rate, relative risks will move towards unity with increasing length of follow up. In large trials, follow up is often longer than in small studies. Finally, an asymmetrical funnel plot may arise by chance.

The trials displayed in a funnel plot may not estimate the same underlying effect of the intervention, and such heterogeneity between results may lead to asymmetry in funnel plots. For example, if a combined outcome is considered then substantial benefit may be seen only in patients at high risk for the component of the combined outcome that is affected by the intervention.[42] A cholesterol lowering drug that reduces mortality from coronary heart disease will have a greater effect on all cause mortality in high risk patients with established cardiovascular disease than in asymptomatic patients with isolated hypercholesterolaemia. This is because a consistent relative reduction in mortality from coronary heart disease will translate into a greater relative reduction in all cause mortality in high risk patients, in whom a greater proportion of all deaths will be from coronary heart disease. This will produce asymmetry in funnel plots if the smaller trials were performed in high risk patients.

Small trials are generally conducted before larger trials are established. In the intervening years, control treatments may have improved or changed in a way that could reduce the efficacy of the experimental treatment. Such a mechanism has been proposed as an explanation for the discrepant results obtained in clinical trials of the effect of magnesium infusion in myocardial infarction,[43] although this interpretation is not supported by the data from clinical trials.[44] Finally, some interventions may have been implemented less thoroughly in larger trials, thus explaining the more positive results in smaller trials. This could have occurred in one of the interventions considered in our comparison of meta-analysis and single large trials, inpatient geriatric consultation.[14]

Very different mechanisms can thus lead to asymmetry in funnel plots, as summarised in the box. It is important to note, however, that this will always be associated with a biased overall estimate of effect when studies are combined in a meta-analysis. The more pronounced the asymmetry, the more likely it is that the amount of bias will be substantial. The exception to this rule arises when asymmetry is produced by chance alone.

How frequent is bias in meta-analysis?

Several studies have recently tried to evaluate the validity of meta-analysis. Villar et al analysed 38 meta-analyses from the pregnancy and childbirth module of the 1993 Cochrane database by comparing the results from the largest trial with the remaining smaller studies.[45] On the basis of the direction of estimates of treatment effects, they concluded that 80% of meta-analyses were in total or partial agreement with the results from the larger "gold standard" trial. In a similar study, Cappelleri et al analysed 79 meta-analyses and concluded that there was agreement between smaller trials and large trials in over 80%.[13] In both these analyses, however, the precision of the large trials was low in a sizeable proportion of comparisons. The larger trials in fact often provided an estimate of lower precision than the meta-analysis of the smaller studies. In this situation, concordance between the two could simply be due to the fact that estimates with large, overlapping confidence intervals are unlikely to be classified as discordant.[46]

We thought that stringent criteria were necessary for identifying single large trials that could sensibly be used to assess the results from meta-analyses of smaller trials. As a result, the large trials used in our analysis on average provided an estimate of considerably greater precision that the corresponding meta-analyses. Despite an extensive literature search, we identified only eight such pairs. The matched pair approach may therefore not be suitable assessing the frequency of misleading meta-analysis. However, our results indicate that an asymmetrical funnel plot makes bias likely. The prevalence of funnel plot asymmetry may thus provide a useful proxy measure to examine the prevalence of biased analyses in the literature. Our findings indicate that bias may be present in a small proportion of meta-analyses published in the *Cochrane Database of Systematic Reviews*. Bias may be considerably more prevalent, however, among meta-analyses published in leading general medicine journals. Whether such bias is likely to affect the conclusions of a systematic review or meta-analysis must be carefully assessed for each case.

Begg and Mazumbar proposed a rank correlation test to measure asymmetry in funnel plots.[47] The method is based on the degree of association between the size of effect estimates and their variances. If publication bias is present, the smaller studies will show the larger effects. A positive correlation between effect size and variance emerges in this situation because the variance of the estimates from smaller studies will also be large. When we applied their test to the eight meta-analyses, it indicated significant ($P < 0.1$) asymmetry for only one meta-analysis (inpatient geriatric consultation[14]). This

Key messages

- Systematic reviews of randomised trials are the best strategy for appraising evidence; however, the findings of some meta-analyses were later contradicted by large trials
- Funnel plots, plots of the trials' effect estimates against sample size, are skewed and asymmetrical in the presence of publication bias and other biases
- Funnel plot asymmetry, measured by regression analysis, predicts discordance of results when meta-analyses are compared with single large trials
- Funnel plot asymmetry was found in 38% of meta-analyses published in leading general medicine journals and in 13% of reviews from the *Cochrane Database of Systematic Reviews*
- Critical examination of systematic reviews for publication and related biases should be considered a routine procedure

indicates that the linear regression approach may be more powerful than the rank correlation test.

Conclusions

In the absence of large, conclusive trials for most medical interventions, systematic reviews based on randomised controlled trials are clearly the best strategy for appraising the evidence. Selection bias and other biases pose a serious threat to the validity of this approach, however, and care must be taken to avoid meta-analysis becoming discredited. The technique discussed here should contribute to this goal, providing a reproducible measure for the likely presence, or apparent absence, of such biases. It is easily calculated and provides summary statistics that can be reported when space limitations do not permit the display of funnel plots. Though more methodological research is required, the critical examination for the presence of publication and related biases should be considered a routine procedure. The capacity to unearth such bias will, however, be limited when meta-analyses are based exclusively on small trials. There is no statistical solution in this situation, and the results from such analyses should therefore be treated with caution.

We are grateful to Andreas Stuck and Gilbert Ramirez for kindly providing additional data.

Funding: Swiss National Science Foundation (grants 3200-045597 and 3233-038803).

Conflict of interest: None.

1 Chalmers I, Dickersin K, Chalmers TC. Getting to grips with Archie Cochrane's agenda. *BMJ* 1992;305:786-8.
2 Mulrow CD. Rationale for systematic reviews. *BMJ* 1994;309:597-9.
3 Huque MF. Experiences with meta-analysis in NDA submissions. *Proc Biopharmaceutical Section Am Statist Assoc* 1988;2:28-33.
4 Egger M, Davey Smith G. Misleading meta-analysis. Lessons from "an effective, safe, simple" intervention that wasn't. *BMJ* 1995;310:752-4.
5 Eysenck HJ. An exercise in mega-silliness. *Am Psychol* 1978;33:517.
6 Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991;337:867-72.
7 Gøtzsche PC. Reference bias in reports of drug trials. *BMJ* 1987;295:654-6.
8 Egger M, Zellweger-Zähner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. *Lancet* 1997;350:326-9.
9 Egger M, Davey Smith G. Meta-analysis: bias in location and selection of studies. *BMJ* (in press).
10 Light RJ, Pillemer DB. *Summing up. The science of reviewing research.* Cambridge, MA: Harvard University Press, 1984.
11 Villar J, Piaggio G, Carroli G, Donner A. Factors affecting the comparability of meta-analyses and largest trials results in perinatology. *J Clin Epidemiol* 1997;50:997-1002.
12 Galbraith R. A note on graphical presentation of estimated odds ratios from several clinical trials. *Stat Med* 1988;7:889-94.
13 Cappelleri JC, Ioannidis JPA, Schmid CH, de Ferranti SD, Aubert M, Chalmers TC, et al. Large trials vs meta-analysis of smaller trials. How do their results compare? *JAMA* 1996;276:1332-8.
14 Stuck AE, Siu AL, Wieland GD, Adams J, Rubenstein LZ. Comprehensive geriatric assessment: a meta-analysis of controlled trials. *Lancet* 1993; 342:1032-6.
15 SOLVD investigators. Effect of enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure. *N Engl J Med* 1991;325:293-302.
16 *The Cochrane Database of Systematic Reviews*. Oxford: Cochrane Collaboration, 1996.
17 Yusuf S, Collins R, Peto R, Furberg C, Stampfer MJ, Goldhaber SZ, et al. Intravenous and intracoronary fibrinolytic therapy in acute myocardial infarction: overview of results on mortality, reinfarction and side-effects from 33 randomized controlled trials. *Eur Heart J* 1985;6:556-85.
18 Gruppo Italiano per lo Studio della Streptochinasi nell'Infarto Miocardico (GISSI). Effectiveness of intravenous thrombolytic treatment in acute myocardial infarction. *Lancet* 1986;i:397-402.
19 Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Progr Cardiovasc Dis* 1985;17:335-71.
20 ISIS-1 Collaborative Group. Randomised trial of intravenous atenolol among 16 027 cases of suspected acute myocardial infarction: ISIS-1. *Lancet* 1986;ii:57-66.
21 Wang PH, Lau J, Chalmers TC. Meta-analysis of effects of intensive blood-glucose control on late complications of type I diabetes. *Lancet* 1993;341:1306-9.
22 Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med* 1993;329:977-86.
23 Reuben DB, Borok GM, Wolde-Tsadik G, Ershoff DH, Fishman LK, Ambrosini VL, et al. Randomized trial of comprehensive geriatric assessment in the care of hospitalized patients. *N Engl J Med* 1995;332:1345-50.
24 Yusuf S, Collins R, MacMahon S, Peto R. Effect of intravenous nitrates on mortality in acute myocardial infarction: an overview of the randomised trials. *Lancet* 1988;i:1088-92.
25 Gruppo Italiano per lo Studio della Streptochinasi nell'Infarto Miocardico (GISSI). GISSI-3: effects of lisinopril and transdermal glyceryl trinitrate singly and together on 6-week mortality and ventricular function after acute myocardial infarction. *Lancet* 1994;343:1115-22.
26 Mulrow CD, Mulrow JP, Linn WD, Aguilar C, Ramirez G. Relative efficacy of vasodilator therapy in chronic congestive heart failure. *JAMA* 1988;259:3422-6.
27 Teo KK, Yusuf S. Role of magnesium in reducing mortality in acute myocardial infarction. A review of the evidence. *Drugs* 1993;46:347-59.
28 ISIS-4 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;345:669-87.
29 Imperiale TF, Stollenwerk Petrullis A. A meta-analysis of low-dose aspirin for the prevention of pregnancy-induced hypertensive disease. *JAMA* 1991;266:261-5.
30 CLASP Collaborative Group. CLASP: a randomized trial of low-dose aspirin for the prevention and treatment of pre-eclampsia among 9364 pregnant women. *Lancet* 1994;343:619-29.
31 Garg R, Yusuf S for the Collaborative Group on ACE Inhibitor Trials. Overview of randomised trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure. *JAMA* 1995;273:1450-6.
32 Collins R. Antiplatelet agents for IUGR and pre-eclampsia. In: Enkin MW, Keirse MJNC, Renfrew MJ, Neilson JP, eds. *Pregnancy and childbirth module, Cochrane Database of Systematic Reviews.* Oxford: Update Software, 1994. (Review No 04000, 12 March 1994. Cochrane Updates on Disk, disk issue 3.)
33 Chalmers I. Underreporting research is scientific misconduct. *JAMA* 1990;263:1405-8.
34 Levy G. Publication bias: its implications for clinical pharmacology. *Clin Pharmacol Ther* 1992;52:115-9.
35 Savulescu J, Chalmers I, Blunt J. Are research ethics committees behaving unethically? Some suggestions for improving performances and accountability. *BMJ* 1996;313:1390-3.
36 Ravnskov U. Cholesterol lowering trials in coronary heart disease: frequency of citation and outcome. *BMJ* 1992;305:15-9.
37 Huston P, Moher D. Redundancy, disaggregation, and the integrity of medical research. *Lancet* 1996;347:1024-6.
38 Chalmers TC, Celano P, Sacks HS, Smith H. Bias in treatment assignment in controlled clinical trials. *N Engl J Med* 1983;309:1358-61.
39 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.
40 Altman DG. The scandal of poor medical research. *BMJ* 1994;308:283-4.
41 Egger M, Davey Smith G, Phillips AN. Meta-analysis: principles and procedures. *BMJ* (in press).
42 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.
43 Baxter GF, Sumeray MS, Walker JM. Infarct size and magnesium: insights into LIMIT-2 and ISIS-4 from experimental studies. *Lancet* 1996;348: 1424-6.
44 Collins R, Peto R. Magnesium in acute myocardial infarction. *Lancet* 1997;349:282.
45 Villar J, Carroli G, Belizan JM. Predictive ability of meta-analyses of randomised controlled trials. *Lancet* 1995;345:772-6.
46 Flournoy N, Olkin I. Do small trials square with large ones? *Lancet* 1995;345:741-2.
47 Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994;50:1088-99.

(Accepted 26 August 1997)