

RESEARCH

Can trial quality be reliably assessed from published reports of cancer trials: evaluation of risk of bias assessments in systematic reviews

 OPEN ACCESS

Claire L Vale *senior research scientist*, Jayne F Tierney *senior research scientist*, Sarah Burdett *senior research scientist*

Meta-analysis Group, MRC Clinical Trials Unit, London WC2B 6NH, UK

Abstract

Objective To evaluate the reliability of risk of bias assessments based on published trial reports, for determining trial inclusion in meta-analyses.

Design Reliability evaluation of risk of bias assessments.

Data sources 13 published individual participant data (IPD) meta-analyses in cancer were used to source 95 randomised controlled trials.

Review methods Risk of bias was assessed using the Cochrane risk of bias tool (RevMan5.1) and accompanying guidance. Assessments were made for individual risk of bias domains and overall for each trial, using information from either trial reports alone or trial reports with additional information collected for IPD meta-analyses. Percentage agreements were calculated for individual domains and overall (<66%=low, ≥66%=fair, ≥90%=good). The two approaches were considered similarly reliable only when agreement was good.

Results Percentage agreement between the two methods for sequence generation and incomplete outcome data was fair (69.5% (95% confidence interval 60.2% to 78.7%) and 80.0% (72.0% to 88.0%), respectively). However, percentage agreement was low for allocation concealment, selective outcome reporting, and overall risk of bias (48.4% (38.4% to 58.5%), 42.1% (32.2% to 52.0%), and 54.7% (44.7% to 64.7%), respectively). Supplementary information reduced the proportion of unclear assessments for all individual domains, consequently increasing the number of trials assessed as low risk of bias (and therefore available for inclusion in meta-analyses) from 23 (23%) based on publications alone to 66 (66%) based on publications with additional information.

Conclusions Using cancer trial publications alone to assess risk of bias could be unreliable; thus, reviewers should be cautious about using them as a basis for trial inclusion, particularly for those trials assessed as unclear risk. Supplementary information from trialists should be sought to enable appropriate assessments and potentially reduce or overcome some risks of bias. Furthermore, guidance should ensure clarity on what constitutes risk of bias, particularly for the more subjective domains.

Introduction

The quality of the studies that contribute to a systematic review will to some extent determine the validity and reliability of the results of that review. Various tools are available to assess the quality of trials, often in terms of allocation concealment and blinding.^{1 2} Indeed, assessments of trial quality or risk of bias have become a common feature of systematic reviews, being a requirement of publication in peer reviewed journals adhering to PRISMA guidelines.³

The Cochrane Handbook⁴ states that because the ability to measure the true bias (or even the true risk of bias) is limited, then the possibility to validate a tool to assess that risk is also limited. Nevertheless, authors of Cochrane systematic reviews are required to use the Cochrane risk of bias tool (RevMan5.1)⁵ to appraise risk of bias for randomised controlled trials across six domains relating to selection, performance, detection, attrition, and outcome reporting biases, and to combine these assessments to evaluate the risk of bias for individual trials. The current version of the Cochrane handbook⁴ also suggests that reviewers should “take risk of bias into account” in meta-analyses. One recommendation is that studies at “low” and “unclear” risk of bias should not be combined in meta-analyses, unless authors “provide specific reasons for believing that these studies are likely to have been conducted in a manner that avoided bias.” Alternatively, studies judged to be at high or unclear risk of bias could be “given reduced weight in meta-analyses,” compared with studies at low risk of bias.⁶ However, since formal statistical methods are not yet recommended for routine use, authors are guided to “restrict meta-analyses to studies at low (or lower) risk of bias, or to stratify studies according to risk of bias.”

The vast majority of Cochrane (and other) systematic reviews is based on information extracted from the publications of eligible studies. Therefore, most risk of bias assessments are similarly based on trial publications. However, trial quality is

not necessarily well represented in publications.⁷ While initiatives such as CONSORT^{8,9} should improve quality, many trials included in systematic reviews may have been published before the uptake of CONSORT guidelines; been published in journals that are yet to implement CONSORT recommendations; or, indeed, not been published at all.

By conducting systematic reviews based on individual participant data (IPD), as well as collating full datasets for each of the included trials, our group has obtained copies of trial protocols and other information about the design and conduct of included trials directly from trial investigators, rather than relying solely on published information. This approach has allowed a greater insight into any potential biases in those trials. We therefore aimed to use this additional information to evaluate the reliability of risk of bias assessments based on trial publications alone for trials included in our meta-analysis. A further aim was to investigate the effect of any differences in the risk of bias judgments for individual trials on resulting meta-analyses.

Methods

Thirteen completed IPD meta-analyses of treatments for cancer, published by our group, were used as a source of randomised controlled trials.¹⁰⁻²¹ So that risk of bias assessments could be conducted, trials had to be published either in full or as an abstract, and a copy of the trial protocol or forms detailing trial design completed by trialists (or both) had to be available. Therefore, unpublished trials and those for which we had neither protocols nor forms were necessarily excluded.

Blinding of the treatment allocations is rarely feasible for randomised controlled trials of cancer treatments, and because the primary outcome of the studies (and indeed for all of the included IPD reviews) is often overall survival, blinded outcome assessments are uncommon and unlikely to introduce bias. We therefore applied an assessment of low risk of bias to the domains of blinding of participants and personnel and blinding of outcome assessment for all included trials. Two authors (SB and CLV) carried out assessments of risk of bias for the individual domains relating to allocation concealment, sequence generation, incomplete outcome data, and selective reporting of outcomes using the Cochrane risk of bias tool⁵ and guidance from the Cochrane handbook.⁴

Data relating to individual domains were first extracted from the trial publications so that assessments of risk of bias could be made. The process was then repeated using the publications plus additional information collected as part of the IPD process (table 1⇓). Because our IPD meta-analyses spanned a period of almost 20 years, we had not designed data collection forms with risk of bias assessment in mind; subsequently, information relating to the individual domains was not consistently requested. Therefore, we sought details directly from the IPD supplied on the numbers of patients randomised or analysed for assessment of attrition bias and on the outcomes available, to establish selective outcome reporting bias. If assessments of any of the individual domains by the two authors disagreed, they were resolved by discussion and consensus, sometimes involving a third author (JFT), in order to obtain a single set of assessments for each trial for the two approaches.

To obtain an overall risk of bias assessment for each trial, the authors agreed a priori on three key domains: sequence generation, allocation concealment, and incomplete outcome data. These domains were thought most likely to represent potential biases within trials. For a trial to be classified as low risk of bias, all three key domains had to be judged low risk. If

one or more of these domains was classed as unclear, the overall judgement for the trial was also unclear; similarly, if one or more of the domains was assessed as high risk of bias, the trial was also deemed to be at high risk of bias.

To assess the reliability of basing assessments on publications alone, we calculated percentage agreements and associated 95% confidence intervals between assessments based on the two approaches for each individual domain and at the trial level. Agreement of less than 66% was considered to be low, whereas agreement of 66% or more was categorised as fair.²² However, for the two approaches to be considered similarly reliable, a high level of agreement, in the order of 90%, was regarded as appropriate.

Finally, we explored the potential effect of the trial level assessments on meta-analyses by comparing the number and proportion of trials assessed as low risk of bias and therefore considered appropriate for inclusion in each of the 13 IPD meta-analyses using the two approaches.

Results

We found 95 randomised controlled trials from 13 completed and published IPD meta-analyses for which publications, protocols, or completed forms were available. Other trials included in these 13 meta-analyses were not available for this study, either because they were unpublished, or because forms or protocols had not been collected (fig 1⇓). Of 95 available trials, 88 (93%) were published in full and seven (7%) had been presented as conference abstracts. Risk of bias assessments, using both approaches, were completed for all 95 studies.

Risk of bias assessments for individual domains

Selection bias

Based on published information alone, 42 (44%) trials were judged at low risk of bias for sequence generation compared with 69 (73%) when additional information from protocols and forms was used. This difference was largely due to a reduction in the number of trials classified as unclear risk of bias (53 (56%) from publications alone decreasing to 26 (27%) when additional information was used). The percentage agreement between the two approaches was fair (69.5%, 95% confidence interval 60.2% to 78.7%; table 2⇓, fig 2⇓).

For allocation concealment, 40 (42%) trials were assessed as having low risk of bias using publications only compared with 89 (93%) when using additional information. Again, this was due to a reduction in the proportion of unclear classifications, from 55 trials (58%) based on publications alone to only six trials (7%) based on the additional information. The percentage agreement between the two approaches was low (48.4%, 95% confidence interval 38.4% to 58.5%; table 2, fig 3⇓). There were no assessments of high risk of bias for either of these domains using either approach.

Performance and detection bias

Although none of the included trials was blinded, for the primary outcome of overall survival, we assessed both performance and detection bias as low risk for all trials and using both approaches for reasons outlined above.

Attrition bias

To evaluate attrition bias, on the basis of whether the outcome data were incomplete or not, the authors had to establish a rule

of thumb to ensure consistency between assessments. Trials were assessed as low risk of bias if less than 10% of patients were excluded overall and if similar proportions were excluded from both arms. Trials were judged as high risk of bias if there were considerable imbalances between arms or if more than 10% of randomised patients were excluded from the analysis.

Based on publications only, 74 trials (78%) were assessed as low risk, compared with 90 (95%) using the additional information. Eleven trials (12%) were judged to be at unclear risk based on publications alone compared with only one trial (1%) when the additional information was used. Ten further trials (11%) were at high risk of attrition bias from the publications alone compared with only four trials (4%) using additional information. Overall, the percentage agreement between the two approaches was fair (80%, 95% confidence interval 72.0% to 88.0%; table 2, fig 4).

Outcome reporting bias

Based on the publications only, 37 trials (39%) were judged to be at low risk of outcome reporting bias compared with 90 trials (95%) when protocols and forms were used. Ten trials (11%) were assessed as unclear risk of bias based on publications alone whereas with additional information, no trials were judged unclear. The number of trials at high risk of bias also fell from 48 (51%) based on publications to five (5%) with additional information. The percentage agreement between the two approaches was low (42.1%, 95% confidence interval 32.2% to 52.0%; table 2, fig 5).

Overall risk of bias assessments for individual trials

Based on publications only, 23 trials (24%) were classified as low risk of bias compared with 64 trials (67%) based on publications supplemented with additional information. This was largely due to the reduction in trials classified as unclear risk of bias from 70 (74%) using publications alone to 31 (33%) using protocols and forms. There were no trials at high risk of bias with the use of additional information compared with two trials (2%) with publications only. The percentage agreement between the two approaches to judging overall risk of bias was low (54.7%, 95% confidence interval 44.7% to 64.7%; table 2, fig 6).

Just over half the trials (n=54, 57%) were published up to or including 1996, when the original CONSORT statement was published,⁸ with the remainder (n=41, 43%) published from 1997 onwards. We found a reduction in the proportion of unclear risk of bias assessments ($\chi^2=5.52$, $P=0.02$) for trials published after CONSORT (≥ 1997 , n=25, 61%) compared with those published before CONSORT (≤ 1996 , n=45, 83%).

Potential impact on meta-analyses

Overall, 23 trials (23%) were assessed as low risk of bias based on publications alone; however, with additional information, 66 trials (66%) were classified as low risk of bias (table 3). Had the 13 meta-analyses included only those trials at low (or lower) risk of bias (as recommended in the Cochrane Handbook), and if assessments were based on publications alone, five meta-analyses (38%) could not be undertaken, because none of the included trials were judged to be at low risk of bias. If additional information was used, the number of trials assessed at low risk of bias (and therefore available for inclusion) was increased for all except for one meta-analysis (table 3), largely because of better ascertainment of sequence generation (fig 2) and allocation concealment (fig 3) from trial

protocols or forms. Therefore, the additional information clearly would have improved the power, precision, and reliability of the results obtained in all but one of the meta-analyses. For example, in one meta-analysis in our sample, the inclusion of only trials judged as having low risk of bias from publications limited the meta-analysis to only two of 14 eligible trials (338 of a total of 3995 randomised patients), and—perhaps not surprisingly—results based on this limited subset of the trials were inconclusive.

Discussion

By comparing assessments of risk of bias of randomised controlled trials in cancer made from publications with those using supplementary information, including IPD, our study has gone some way towards validating the Cochrane risk of bias tool, since access to that additional information and data enables us to get closer to the true risk of bias of individual studies. Appraising trial quality should inevitably be a key aspect of any well conducted systematic review; however, our results indicate that basing such assessments on publications alone is probably not appropriate. In general, agreement between the two approaches was low, and for all but one of the individual domains (attrition bias) and the overall assessment for the trial, agreement fell far short of “ideal.”

Comparison with other studies

These findings are supported by a recent study that showed an increased proportion of “adequate” quality assessments for sequence generation and allocation concealment in 429 randomised controlled trials in cancer when trial protocols were used alongside publications.²³ A further study found that the assessment of selective outcome reporting bias was changed for three of five trials for which protocols were obtained in a systematic review of treatments for asthma.²⁴ Our results show that when all information was taken into account, the majority of trials assessed as unclear risk of bias from publications alone were actually at low risk of bias. Therefore, current advice regarding trials at unclear risk of bias might not be appropriate. Our study indicates that deficiencies with the reporting of trials does not necessarily reflect deficiencies in trial quality, and that poor reporting can often be the cause of inappropriate evaluations of risk of bias, in particular for selective outcome reporting bias.

Even for trials from the post CONSORT⁸ era, more than 60% were deemed to have unclear risk of bias using publications alone. Although wider adoption of CONSORT and greater implementation of its requirements by journal editors, peer reviewers, and trial authors might improve the ability to judge a trial's quality from its publication, some trials will inevitably remain unreported, or be reported only in the grey literature. Although in due course, CONSORT for abstracts²⁵ could go some way to improving this reporting situation, risk of bias assessments for unpublished trials remain difficult. Their inclusion in systematic reviews, however, remains fundamental to reducing the effect of reporting biases.²⁶⁻²⁹

Strengths and limitations of study

Assessing risk of bias was particularly difficult for the more subjective domains. Although the two authors made consistent judgments for allocation concealment and sequence generation, discrepancies were more common for attrition bias such that a rule around cut-off rates of attrition or patient exclusions was established by the authors to ensure consistency of assessments. Clearly, a different cut-off would lead to different results.

Consistent assessment of selective outcome reporting bias^{30,31} was also problematic, possibly because it requires reviewers to consider how the reporting of trial level outcomes affects the review as a whole, and is not solely a within trial judgment, as for the other domains. Our findings are similar to those previously reported,³² in which low inter-rater agreements were identified for individual domains, in particular, incomplete outcome data and selective outcome reporting. The authors subsequently demonstrated improved agreement, partly because of their use of specifically developed decision rules for completing risk of bias assessments.²⁴ The 95 randomised controlled trials included in the present study clearly represented a selected group of trials. All of the included trials were cancer trials that, in general, tended to be fairly well conducted. Also, we applied the risk of bias assessment to overall survival—an objective outcome that is commonly well reported—rather than considering all possible outcomes as is recommended. Our results could have therefore represented an optimistic view of the reliability of the risk of bias assessments using published information alone, particularly in relation to incomplete outcome data and selective outcome reporting. Furthermore, the additional information supplied was sometimes limited because the data collection forms obtained in some of the older IPD meta-analyses did not request specific details on methods of sequence generation and planned trial outcomes. The older trial protocols were also often ambiguous with regard to the assessment of some of the required domains. Therefore, even with additional information, around a third of the included studies were classified as unclear risk of bias. Clearly, forms purposely designed to collect specific information would help reviewers reach appropriate judgments regarding risk of bias, in particular for those trials with inadequate information published.

In an IPD meta-analysis, we routinely checked the integrity of the randomisation and allocation concealment for the IPD supplied³³ such that, in reality, none of the 95 trials included in this study was excluded from our meta-analyses for being at unclear or high risk of these selection biases. If uncertainty about trial design or conduct could not be adequately resolved—for example, if many patients were excluded from the trial and we were not able to reinstate them, meaning that the trial potentially had a high risk of attrition bias—we would report this and conduct sensitivity analyses, thus using risk of bias to more fittingly influence the meta-analyses. However, this approach depended on obtaining a reliable assessment of risk of bias.

Conclusions

Our results have shown that obtaining additional information about allocation concealment, sequence generation methods, and blinding can improve the accurate ascertainment of selection, performance, and detection biases. This approach can therefore ensure that trials are not inappropriately excluded from meta-analyses, simply because of inadequate reporting. Information can be sought directly from trialists, trial protocols, or trial registries.³⁴ If resources are limited, the greatest attention could be given to those trials with the most limited information, such that appropriate assessments can be made for all eligible trials. Furthermore, obtaining information regarding numbers of patients randomised or planned outcome assessments may also overcome deficiencies in the reporting of trials. Indeed, if summary results are obtained, it may also be possible to overcome the risk of attrition or selective outcome reporting bias completely within a meta-analysis.

Guidance needs to be clear on what constitutes risk of bias, particularly for the more subjective bias domains. Advice to exclude trials at unclear risk of bias could be misleading. Certainly, reviewers should be cautious about basing decisions about trial inclusion on the risk of bias, particularly if assessments have been obtained using publications alone, which may lead to good evidence being disregarded in meta-analyses.

We thank the trialists who have previously supplied copies of trial protocols and forms for inclusion in IPD meta-analyses cited in this article.

Contributors: All authors were involved in the design and conduct of the analyses and interpretation of the results. CLV drafted the manuscript, which was revised by SB and JFT. All authors have read and agreed the final manuscript. CLV is the study guarantor.

Funding: This work was supported by the United Kingdom's Medical Research Council.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/doi_disclosure.pdf and declare: support from the Medical Research Council for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Data sharing: No additional data available.

- Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary. *Control Clin Trials* 1996;17:1-12.
- 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.
- Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009;339:b2700.
- Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions* 4.2.5 [updated May 2005]. Vol Issue 3. John Wiley and Sons, 2005.
- Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.
- Spiegelhalter DJ, Best NG. Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Stat Med* 2003;22:3687-709.
- Soares HP, Daniels S, Kumar A, Clarke M, Scott C, Swasn S, et al. Bad reporting does not mean bad methods for randomised trials: observational study of randomised controlled trials performed by the Radiation Therapy Oncology Group. *BMJ* 2004;328:22-5.
- Altman DG. Better reporting of randomised controlled trials: the CONSORT statement. *BMJ* 1996;313:570-1.
- Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663-94.
- Non-small Cell Lung Cancer Collaborative Group. Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials. *BMJ* 1995;311:899-909.
- Sarcoma Meta-analysis Collaboration. Adjuvant chemotherapy for localised resectable soft tissue sarcoma in adults: meta-analysis of individual patient data. *Lancet* 1997;350:1647-54.
- PORT Meta-analysis Trialists Group. Postoperative radiotherapy in non-small-cell lung cancer: systematic review and meta-analysis of individual patient data from nine randomised controlled trials. *Lancet* 1998;352: 257-63.
- Non-small Cell Lung Cancer Collaborative Group. Chemotherapy for non-small cell lung cancer. *Cochrane Database Syst Rev* 2000;2:CD002139.
- Glioma Meta-analysis Trialists Group. Chemotherapy for high-grade glioma. *Cochrane Database Syst Rev* 2002;3:CD003913.
- Neoadjuvant Chemotherapy for Cervix Cancer Meta-analysis (NACCMA) Collaboration. Neoadjuvant chemotherapy for locally advanced cervical cancer: a systematic review and meta-analysis of individual patient data from 21 randomised trials. *Eur J Cancer* 2003;39:2470-86.
- Arnott SJ, Duncan W, Gignoux M, David GJ, Hansen HS, Launois B, et al. Preoperative radiotherapy in esophageal carcinoma: a meta-analysis using individual patient data (Oesophageal Cancer Collaborative Group). *Int J Radiat Oncol Biol Phys* 1998;41:579-83.
- Advanced Bladder Cancer (ABC) Meta-analysis Collaboration. Neoadjuvant chemotherapy in invasive bladder cancer: a systematic review and meta-analysis. *Lancet* 2003;361:1927-34.
- Advanced Bladder Cancer (ABC) Meta-analysis Collaboration. Adjuvant chemotherapy in invasive bladder cancer: a systematic review and meta-analysis of individual patient data. *Eur Urol* 2005;48:189-201.
- Chemoradiotherapy for Cervical Cancer Meta-analysis Collaboration. Reducing uncertainties about the effects of chemoradiotherapy for cervical cancer: a systematic review and meta-analysis of individual patient data from 18 randomized trials. *J Clin Oncol* 2008;26:5802-12.
- NSCLC Meta-analyses Collaborative Group. Chemotherapy in addition to supportive care improves survival in advanced non-small-cell lung cancer: a systematic review and meta-analysis of individual patient data from 16 randomized controlled trials. *J Clin Oncol* 2008;26:4617-25.

What is already known on this topic

Flawed design, conduct, analysis, and reporting can lead to poor treatment evaluations in randomised trials; however, poor trial reporting does not necessarily reflect poor methodological quality of the trial

The Cochrane Collaboration's risk of bias tool, used to assess potential for bias in randomised trials across six domains, is now mandatory for Cochrane systematic reviews of randomised controlled trials

Risk of bias assessment of trials included in systematic reviews is commonly based on published reports and has been suggested as a way to determine the inclusion of trials, or weighting of their results, in meta-analyses

What this study adds

Access to additional information from randomised controlled trials in cancer has enabled validation of the Cochrane risk of bias tool for assessing reported trial quality

Risk of bias assessments from published trial reports alone could be unreliable such that reviewers should be cautious about using them as a basis for trial inclusion in meta-analyses; use of additional information or summary statistics, if available, could reduce or overcome potential biases

Guidance should ensure clarity on what constitutes risk of bias, particularly for domains with more subjective assessments, and should not direct reviewers to exclude potentially good evidence from reviews

- 21 NSCLC Meta-analyses Collaborative Group. Adjuvant chemotherapy, with or without postoperative radiotherapy, in operable non-small cell lung cancer: two meta-analyses of individual patient data. *Lancet* 2010;375:1267-77.
- 22 McMinn AM, van Sluijs EMF, Harvey NC, Cooper C, Inskip HM, Godfrey KM, et al. Validation of a maternal questionnaire on correlates of physical activity in preschool children. *Int J Behav Nutr Phys Act* 2009;6:81.
- 23 Mhaskar R, Djulbegovic B, Magazín A, Soares HP, Kumar L. Published methodological quality of randomized controlled trials does not reflect the actual quality assessed in protocols. *J Clin Epidemiol* 2012;65:602-9.
- 24 Hartling L, Bond K, Vandermeer B, Seida J, Dryden DM, Rowe BH. Applying the risk of bias tool in a systematic review of combination long-acting beta-agonists and inhaled corticosteroids for persistent asthma. *PLoS One* 2011;6:e17242.
- 25 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med* 2008;5:e20.
- 26 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.
- 27 Sutton AJ, Duval SJ, Tweedie RL, Abrams KR, Jones DR. Empirical assessment of effect of publication bias in meta-analyses. *BMJ* 2000;320:1574-7.
- 28 Hopewell S, Clarke M, Stewart L, Tierney J. Time to publication for results of clinical trials. *Cochrane Database Syst Rev* 2007;2:MR000011.
- 29 McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in a meta-analysis. *Lancet* 2000;356:1228-31.
- 30 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.
- 31 Dwan K, Gamble C, Kollmannage-Dona R, Mohammed S, Powell C, Williamson PR. Assessing the potential for outcome reporting bias in a review: a tutorial. *Trials* 2010;11:52.
- 32 Hartling L, Ospina M, Liang Y, Dryden DM, Hooton N, Krebs Seida J, et al. Risk of bias versus quality assessment of randomised controlled trials: cross sectional study. *BMJ* 2009;339:b4012.
- 33 Stewart LA, Tierney JF. To IPD or Not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Eval Health Prof* 2002;25:76-97.
- 34 Viergever RF, Ghersi D. Information on blinding in registered records of clinical trials. *Trials* 2012;13:210.

Accepted: 12 March 2013

Cite this as: *BMJ* 2013;346:f1798

This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 3.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <http://creativecommons.org/licenses/by/3.0/>.

Tables

Table 1 | Sources of information from which data for individual risk of bias domains were obtained

Risk of bias domain/information sought	Sources of information used to extract data relating to risk of bias (where available)			
	Trial publication	Trial protocol	Data collection forms	Individual patient data
Selection bias				
Method of allocation concealment	Yes	Yes	Yes	No
Method of sequence generation	Yes	Yes	Yes	No
Attrition bias				
Number of patients randomised (by arm)	Yes	No	Yes	Yes
Number of patients excluded (by arm)	Yes	No	Yes	Yes
Reasons for exclusion	Yes	No	No	Yes
Selective outcome reporting bias				
Outcomes planned	Yes	Yes	Yes	No
Outcomes reported	Yes	No	Yes	Yes
Outcomes available	Yes	No	Yes	Yes

Table 2| Outcomes and comparison of risk of bias assessments

Risk of bias domain	No of assessments based on publications only			No of assessments based on publications plus supplementary information			Percentage agreement (%; 95% CI)
	Low	Unclear	High	Low	Unclear	High	
Sequence generation	42	53	0	69	26	0	69.5 (60.2 to 78.7)
Allocation concealment	40	55	0	89	6	0	48.4 (38.4 to 58.5)
Incomplete outcome data	74	10	11	90	1	4	80.0 (72.0 to 88.0)
Selective outcome reporting	37	10	48	90	0	5	42.1 (32.2 to 52.0)
Overall risk of bias for trial	23	70	2	64	31	0	54.7 (44.7 to 64.7)

Table 3| Studies assessed as low risk of bias for individual meta-analyses

Meta-analysis	No of trials with completed assessments	Risk of bias assessment					
		Publications only			Publications plus extra information		
		Low	Unclear	High	Low	Unclear	High
1	14	2	12	0	14	0	0
2	8	5	3	0	6	2	0
3	3	0	3	0	2	1	0
4	8	1	6	1	6	2	0
5	2	0	2	0	1	1	0
6	9	3	5	1	6	3	0
7	3	0	3	0	1	2	0
8	8	1	7	0	2	6	0
9	5	2	3	0	2	3	0
10	15	5	10	0	11	4	0
11	17	4	13	0	10	7	0
12	6	0	6	0	4	2	0
13	2	0	2	0	1	1	0
Total	100	23 (23%)	75 (75%)	2 (2%)	66 (66%)	34 (34%)	0

Data are no of assessments.

*Percentage of total number of trials. Total number of trials (n=100) is greater than the total number of unique trials in the sample (n=95), because some trials contributed data to more than one meta-analysis).

Figures

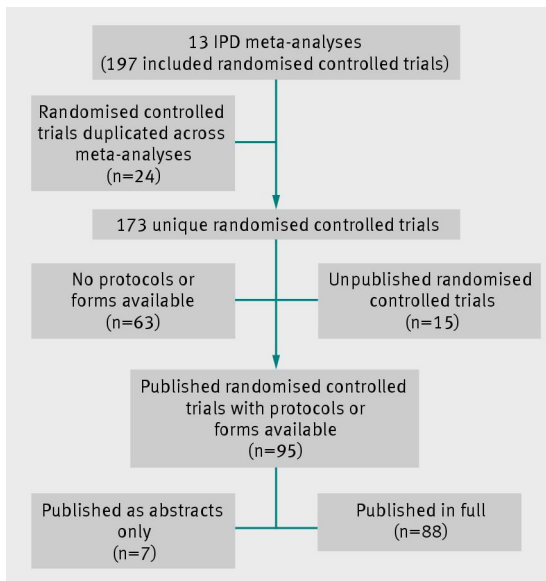


Fig 1 Flowchart of trial sample selection

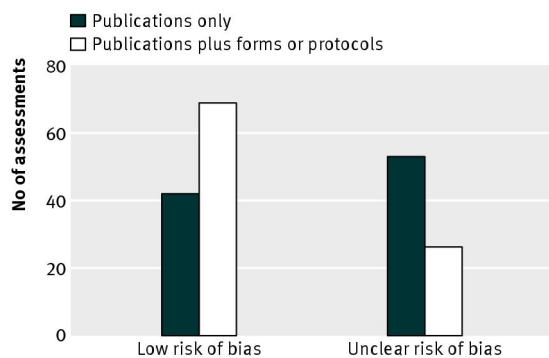


Fig 2 Comparison of assessments for sequence generation between publications only and publications plus additional information

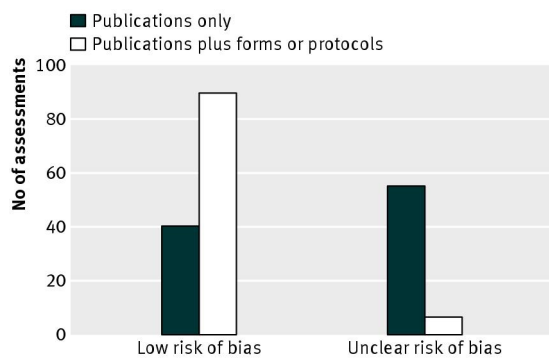


Fig 3 Comparison of assessments for allocation concealment between publications only and publications plus additional information

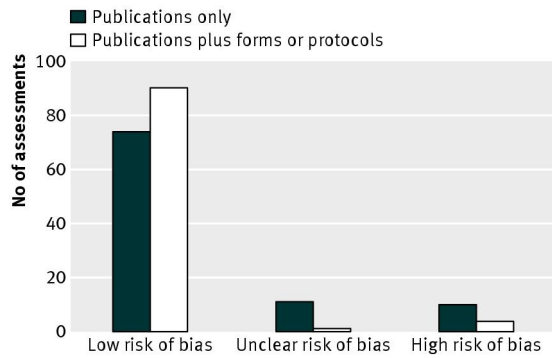


Fig 4 Comparison of assessments for incomplete outcome data between publications only and publications plus additional information

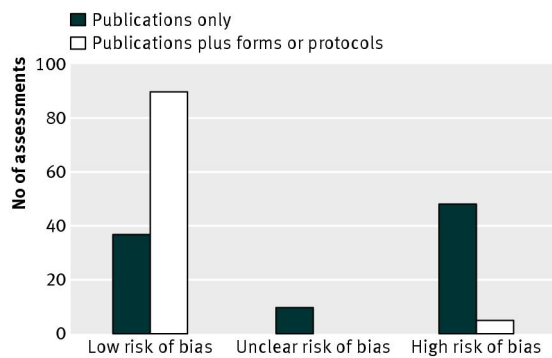


Fig 5 Comparison of assessments for selective outcome reporting between publications only and publications plus additional information

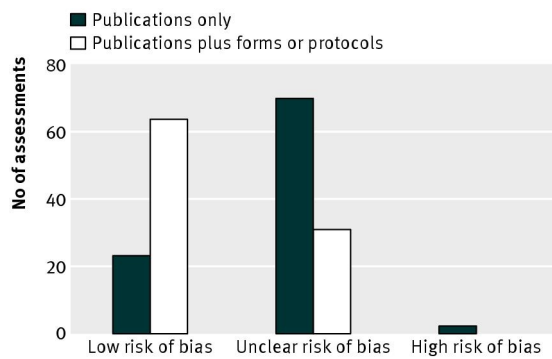


Fig 6 Comparison of assessments for overall risk of bias for trials between publications only and publications plus additional information