ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/5457231

Few systematic reviews exist documenting the extent of bias: a systematic review

Article *in* Journal of Clinical Epidemiology · June 2008

Impact Factor: 3.42 · DOI: 10.1016/j.jclinepi.2007.10.017 · Source: PubMed

CITATIONS
29

READS
34

7 authors, including:

Andrea Tricco
St. Michael's Hospital
155 PUBLICATIONS **2,356** CITATIONS
SEE PROFILE

Margaret J Sampson
Children's Hospital of Eastern Ontario
137 PUBLICATIONS **4,492** CITATIONS
SEE PROFILE

Tanya Horsley
Royal College of Physicians and Surgeons …
52 PUBLICATIONS **1,321** CITATIONS
SEE PROFILE

All in-text references underlined in blue are linked to publications on ResearchGate, letting you access and read them immediately.

Available from: Tanya Horsley
Retrieved on: 08 May 2016

ELSEVIER

Journal of Clinical Epidemiology 61 (2008) 422–434

Journal of Clinical Epidemiology

REVIEW ARTICLES

Few systematic reviews exist documenting the extent of bias: a systematic review

Andrea C. Tricco[a,b], Jennifer Tetzlaff[a,c], Margaret Sampson[a,d], Dean Fergusson[c], Elise Cogo[a,b], Tanya Horsley[c], David Moher[a,c,*]

[a]*Chalmers Research Group, Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Road, Ottawa, Ontario K1N 6N5, Canada*
[b]*Institute of Population Health, University of Ottawa, Ottawa, Ontario, Canada*
[c]*Department of Epidemiology & Community Medicine, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada*
[d]*Department of Information Studies, University of Wales, Aberystwyth, Wales, UK*

Accepted 23 October 2007

Abstract

Objective: To summarize the evidence concerning bias and confounding in conducting systematic reviews (SRs).

Study Design and Setting: Literature was identified through searching the Cochrane Library, MEDLINE, PsycINFO until November 2006, and the authors' files. Studies were included if they were SRs of bias that can occur while conducting a SR. Risk of bias in the SRs was appraised using the Oxman and Guyatt index.

Results: Ten SRs were included. All examined biases related to searching for evidence (e.g., publication bias). One also reported bias associated with obtaining data from included studies (e.g., outcome reporting bias). To minimize bias, data suggest including unpublished material, hand searching for additional material, searching multiple databases, assessing for publication bias, and periodically updating SRs. No SRs were found examining biases related to choosing studies for inclusion or combining studies.

Conclusions: There is little evidence from SRs to support commonly practiced methods for conducting SRs. No SRs summarized studies with prospective designs and most had moderate or minimal risk of bias. Future research should examine bias that can occur during the selection of studies for inclusion and the synthesis of studies, as well as systematically review the existing empirical evidence. © 2008 Elsevier Inc. All rights reserved.

Keywords: Systematic review; Bias; Research methodology; Publication bias; Meta-analysis; Outcome reporting bias

1. Introduction

Systematic reviews (SRs) are becoming increasingly popular in evidence-based health care [1] and have as their strength, methodological features designed to minimize bias. However, evidence suggests that a large proportion of SRs are poorly reported and susceptible to bias [2,3]. For example, a recent cross-sectional survey found that less than half of published SRs reported using a protocol (46.3%) [2]. The use of protocols minimizes bias because hypotheses and methods are stated a priori without prior knowledge of results [4]. Furthermore, the importance of assessing publication bias in SRs has been clearly established [5–7], yet few published SRs consider issues related to publication bias (31.3%) [2]. These findings have serious implications; SRs are often used in the development of clinical practice guidelines [8] and are increasingly viewed as a useful tool for health decision makers [1,9,10].

Bias that can occur while conducting a SR has been described previously. Fifteen years ago in the *journal*, Felson put forth a framework of such biases [11]. This framework explained biases in finding all studies (e.g., publication bias, citation bias); biases that can occur while choosing studies to include in the SR (e.g., inclusion criteria bias, selector bias); and obtaining accurate data bias (e.g., bias in scoring study quality, outcome reporting bias).

Evidence-based information regarding the biases outlined in Felson's framework [11] would provide guidance when conducting a SR. This information can be obtained from SRs that summarize the evidence on biases explained by Felson [11]. We aimed to summarize the evidence for minimizing bias and confounding in conducting SRs and examine gaps in this literature by conducting a SR.

* Corresponding author. Chalmers Research Group, Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Road, Ottawa, Ontario K1N 6N5, Canada. Tel.: 613-738-3591; fax: 613-738-4800.
E-mail address: dmoher@uottawa.ca (D. Moher).

0895-4356/08/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi: 10.1016/j.jclinepi.2007.10.017

What is New?

- To minimize bias during the conduct of systematic reviews (SRs), evidence suggests that the authors should include unpublished material, update SRs periodically, search multiple databases, conduct hand searches, use the Cochrane Highly Sensitive Search Strategy to locate randomized controlled trial reports, and assess for publication bias.
- Additional empirical research examining language bias, outcome reporting bias, the effects of study risk of bias assessment, and the effects of blinding reviewers during a SR have not been summarized by a SR.
- Future research should also examine bias that can occur during the selection of studies for inclusion and bias that occurs during the synthesis of studies.
- SRs examining many widespread practices for conducting SRs were not identified. Based on the identified gaps in the SR literature, perhaps methodological reviewers have fallen behind.

2. Methods

2.1. Eligibility criteria

A SR was defined as any study for which ''the authors' stated objective was to summarize evidence from multiple studies and the article described explicit methods, regardless of the details provided'' [2,12]. When it was clear that the intent of the authors was a literature review (e.g., authors identified the review as a brief overview with no specific review question), as opposed to a SR, articles were excluded [2,12]. We included SRs of empirical studies examining bias and confounding that can occur during the conduct of a SR. The following types of biases relevant to our review identified a priori included publication, indexing/citation, language, time lag, multiple/duplicate publication, outcome reporting, and study quality biases. These were identified from two articles outlining bias in SRs [5,11] and discussions between the investigators.

2.2. Search strategy

SRs were identified through electronic searches in MEDLINE (1966 to November, Week 3 2006, Ovid interface), PsycINFO (1806 to November, Week 4 2006, Ovid interface), the Cochrane Library (2006, Issue 4, Wiley interface), limited to the English language, and the PubMed ''related articles'' link for all references from the Quality of Reporting of Meta-analyses Statement checklist items [13]. Electronic searches were supplemented by using studies from the authors' personal files, contacting experts for gray literature and additional material, and scanning the reference lists of all included SRs. The electronic search strategies were developed and validated by two information specialists (E.C., M.S.; Appendix). Electronic searches were conducted on February 1, 2006 and updated on November 20, 2006.

2.3. Screening

One reviewer (A.C.T.) screened the titles and abstracts identified by the literature search for study inclusion using a predefined study relevance form. This was verified by a second reviewer (J.T.) who screened a random sample of 1/3 of the records. The full text of potentially relevant articles was obtained for further evaluation to determine inclusion.

2.4. Data abstraction

Data were abstracted by one reviewer (A.C.T.), using a prespecified standardized 20-item data abstraction form, and verified by a second reviewer (J.T.) using a 1/5 random sample. Abstracted data included study characteristics (e.g., first author, country or countries where the research originated); the number, study designs, and methodological quality of studies included in the SR; types of bias examined; author's definition of each type of bias; and the SR results. The estimated effect size of bias (e.g., relative risk) and respective confidence intervals (CIs) were also abstracted.

The SR biases were categorized as follows: 1) biases in finding all studies (sampling bias), 2) biases in selecting studies for inclusion, 3) biases in obtaining accurate data from selected studies, and 4) biases that occur when studies are combined [11].

2.5. Risk of bias assessment

The Oxman and Guyatt Overview Quality Assessment Questionnaire was used to assess the risk of bias in the included SRs [14]. This validated instrument consists of nine main criteria for assessing the scientific quality of review articles [14]. The final item asks the assessor to rate the overall scientific quality of the SR using a score ranging from 1 (i.e., extensive flaws) to 7 (i.e., minimal flaws). Two reviewers conducted a training exercise using this instrument (A.C.T., J.T.) and independently rated all studies (A.C.T., J.T.). Disagreements were resolved through discussion.

2.6. Data synthesis

The agreement between the reviewers who screened the literature (A.C.T., J.T.) was assessed using a kappa statistic [15]. We determined a priori that an acceptable level of agreement would be greater than 60% [15]. Results were summarized narratively and quantitative results from the relevant SRs were visually presented in a forest plot.

3. Results

A total of 3,733 records were identified through the searches and subsequently screened. Of these, 221 full-text articles were obtained for further examination to determine relevance, and 10 SRs met our eligibility criteria [5,16–24] (Fig. 1). One of these SRs [23] was identified as an update of a previous SR [16], leaving a total of nine unique relevant SRs. We also identified six Cochrane reviews published as protocols [25–30], which will be included in any subsequent update of this SR (Table 1). Good agreement was observed between the two reviewers (A.C.T., J.T.) at the full-text screening level (kappa = 0.67).

On average, the SRs included 35 studies (range: 2–79) with retrospective cohort, case study, and/or cross-sectional, study designs, and all were published between 2000 and 2005 (Table 2). Half of the reviews reported or conducted quantitative data synthesis (Fig. 2).

Of all included SRs, one did not assess the risk of bias [22] and two did not report assessing the risk of bias [5,24]. The remaining SRs assessed the risk of bias using a component approach [19,23], a checklist [17,18], and by examining other methodological issues (e.g., response rate, databases searched) [20,21] (Table 3).

The risk of bias in the SRs themselves varied greatly. The Oxman and Guyatt index [14] indicated that one SR had major flaws [24], four had minor flaws [5,21–23], and the remaining had minimal flaws [17–20]. Studies were consistently assessed as having an increased risk of bias on the study selection ($n = 6$) and validity ($n = 5$) criteria (Table 4).

3.1. SR biases (Definitions can be found in the glossary of biases.)

3.1.1. Bias in identifying studies (sampling bias)

3.1.1.1. A) Publication-related biases

3.1.1.1.1. Publication bias. Two reviews examining publication bias were identified [5,24]. Seventeen of the 26 (65%) included studies overlapped between both SRs. Song et al. provided information specific to publication bias and included 11 cross-sectional studies [5] (Table 5). The percentage of statistically significant results ranged from 35% to 97% across emergency, medical, biological, and psychology journals [31–35]. The existence of publication bias was consistently confirmed by four retrospective cohort studies of research approved by research ethics boards and trial registries [6,36–38]. The results of a quantitative data synthesis for which a comprehensive literature search was not performed were reported in the Song et al.'s SR [7]. The overall adjusted odds ratio for publication bias was 2.54 (95% CI: 1.44, 4.47; Fig. 2).

Dubben and Beck-Bornholdt included 26 retrospective cohort and cross-sectional studies examining publication bias [24]. The median effect size was 2.3 (CI not reported), indicating that published studies were more than twice as likely to report positive results.

3.1.1.1.2. Gray literature bias. Two SRs examined the effects of including gray literature in SRs [5,18]. There was no overlap in the included studies in these reviews. Song et al. included one retrospective cohort and four cross-sectional studies, and found that trials with statistically positive results were more likely to be published [5]. Hopewell et al. included eight cross-sectional studies [18]. Overall, published trials included more participants (median 46 [interquartile range: 4–300] vs. 5.5 [interquartile range: 4–88]), were more likely to have statistically significant results (30% vs. 19%, $P < 0.05$), and were less susceptible to bias than gray literature trials [18].

3.1.1.1.3. Funding bias. Two SRs provided insight into bias pertaining to source of funding [20,22]. Lexchin et al. included 30 primary studies, 8 (27%) of which were also included in Bekelman et al. On the basis of the 18 comparisons from 15 cohort, cross-sectional, and case studies, pharmaceutical sponsorship was associated with positive outcomes (odds ratio: 4.05, 95% CI: 2.98, 5.51; Fig. 2; Table 5) [20]. All 13 studies examining an association between study quality and funding source found that industry-funded studies were of comparable quality to studies with other funding sources. Bekelman et al. reported similar results in their SR (Fig. 2; Table 5).

3.1.1.1.4. Time-lag bias. Two SRs examined time-lag bias [5,17]. One SR included two retrospective cohort studies, which followed trials from the date of follow-up completion or date of ethics approval to the date of publication [17]. These studies found that 55–58% of all trials were published in full and trials with statistically significant results in favor of the experimental intervention were published quicker (range: 4–5 years) than those with nonsignificant results (range: 6–8 years, $P < 0.05$) [17]. Song et al. also included these two studies, as well as another retrospective cohort study and a cross-sectional study [5]. In the retrospective cohort study included by Song et al., the hazard ratio for time to publication from trial completion for positive vs. negative trials was 3.7 (95% CI: 1.8, 7.7) [39]. In a cross-sectional study of 26 meta-analyses included by Song et al., the treatment effect was exaggerated in most trials published early in the

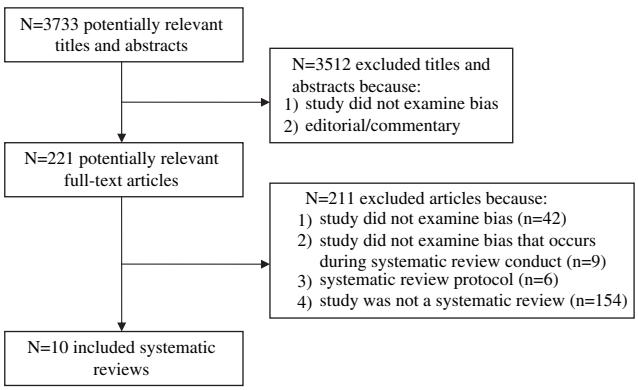

Fig. 1. Study flow chart.

Table 1
Systematic review protocols identified from the search

Author/year, reference	Topic	Search methods	Search limits	Bias category
Clarke/2007 [25]	Individual patient MA vs. aggregate MA	CMR, MEDLINE, EMBASE, hand search, Internet search, contact experts, PubMed-related articles[a]	NR	Combining studies bias
Leeflang/2007 [26]	Identifying diagnostic accuracy studies using MEDLINE and EMBASE searches	MEDLINE, EMBASE, CMR, ISI Web of Science until November 2005, PubMed-related articles, search reference lists	NR	Sampling bias
McDonald/2007 [27]	Identifying randomized trials using MEDLINE searches	MEDLINE (1966 onwards), EMBASE (1980 onwards), CINAHL (1982 onwards) LISA (1969 onwards), SCI (1987 onwards), SIGLE (1980 onwards), hand search, PubMed-related articles, SCI, search reference lists, author contact	NR	Sampling bias
Olsen/2007 [28]	Publication bias in clinical trials	CMR, MEDLINE, contact experts, search reference lists[a]	NR	Sampling bias
Song/2007 [29]	Indirect comparisons	CDSR, DARE, MEDLINE, hand search, search reference lists[a]	NR	Combining studies bias
Westby/2007 [30]	Masking reviewers for study inclusion	CMR, MEDLINE (1966 onwards), EMBASE (1980 onwards), search reference lists, SCI contact experts	NR	Choosing studies bias

Abbreviations: MA, meta-analysis; CMR, Cochrane Methodology Register; NR, not reported; SCI, Science Citation Index; CINAHL, Cumulative Index to Nursing and Allied Health Literature; LISA, Library and Information Science Abstracts; SIGLE, System for Information on Grey Literature in Europe; CDSR, Cochrane Database of Systematic Reviews; DARE, Database of Abstracts of Review of Effects.

[a] Years of search not reported.

drug development cycle compared with later cycle trials (average difference in relative odds: 35%; 95% CI: 15, 55) [40].

3.1.1.1.5. Abstract to full publication bias. Three SRs examined abstract to full publication bias [5,21,23] and findings were consistent across all reviews. Scherer et al. included 79 retrospective cohort studies, 10 (13%) of which were reviewed in the von Elm et al.'s SR ($n = 21$ included studies). Scherer et al. determined that 44.5% of abstracts presented at meetings were subsequently published in full text (95% CI: 43.9, 45.1) [23]. After 9 years, 52.6% of all abstracts, 63.1% of randomized controlled trial (RCT) abstracts, and 49.3% of abstracts with other study designs were published in full. Positive results were associated with full publication; as was oral presentation, acceptance for meeting presentation, RCT design, higher-quality RCT abstracts, and clinical research (vs. basic science research; Fig. 2; Table 5).

von Elm et al. also examined the full publication of results that were presented at meetings [21]. They found similar survival analysis results and noted that abstracts with positive vs. negative outcomes and oral presentations vs. poster presentations were more likely to be published. However, they found that abstracts about basic science were more likely to be published than those on clinical science. Rejected abstracts had the same long-term publication rate as accepted abstracts in this SR (Fig. 2; Table 5).

Song et al. described 22 studies that examined full publication bias [5]. The three studies that were not included by Scherer et al. and von Elm et al. found similar results to those reported above.

3.1.1.1.6. Place of publication bias. One review (Song et al.) summarized three cross-sectional studies that examined place of publication bias. These cross-sectional studies found that the *British Medical Journal* was more likely to publish articles on the "early life hypothesis" (e.g., relationships between indicators of fetal development and later disease patterns) than *The Lancet* [41]. Furthermore, journals considered to be "prominent," such as *Cancer* and the *New England Journal of Medicine* published a higher proportion of positive trials, whereas less well-known journals only published trials with statistically negative results [42].

3.1.1.1.7. Country of conduct bias. Song et al. summarized two cross-sectional and one case study that examined country of conduct bias [5]. In one of these studies, the estimated efficacy of a complementary and alternative therapy was greater in studies published in journals outside the United States when compared to those published in the United States [43]. In another study, the proportion of positive results in trials from China, Taiwan, Japan, and Hong Kong was 100% compared to 56.7% for similar trials published in journals of other countries including Canada, the United States, and Germany [44].

3.1.1.1.8. Language bias. Song et al. summarized five cross-sectional studies that examined language bias and found conflicting evidence for the impact of including or excluding non-English language reports in SRs [5].

3.1.1.2. B) Locating studies using electronic databases

3.1.1.2.1. Indexing bias. Two SRs examined indexing bias [5,19]. Song et al. reviewed eight cross-sectional studies that consistently reported an increased likelihood of missing relevant reports when only one electronic database is searched. Hopewell et al. examined the effects of hand searching vs. electronic searching for identifying reports of RCTs and included 34 cross-sectional studies [19], none of which overlapped with the Song et al.'s SR. Hand searching identified 92–100% of RCTs compared to 42–80% of RCTs when

Table 2
Relevant systematic review characteristics

Author/year, reference	No. of studies included and study designs	Search methods	Search limits	SR bias category
Song/2000 [5]	64 RC, CS, case study	CMR, MEDLINE (1966–1997), EMBASE (1981–1997), BIDS (1981–1997), LISA, PsycLIT (1967–1997), Sociofile (1963–1997), ERIC (1966–1997), DA (1986–1997), MathSci (1940–1997), BEI (1976–1997), SIGLE (1980–1997), ASSIA (1987–1997), search reference lists, contact experts	NR	Sampling bias & obtaining accurate data bias
Hopewell/2001 [17]	2 RC	CMR, MEDLINE, EMBASE, SCI, PubMed-related articles, search reference lists, hand search, contact experts[a]	NR	Sampling bias
Hopewell/2002 [19]	34 CS	CMR (2002), MEDLINE (1966–2002), EMBASE (1980–2002), AMED (1985–2002), BIOSIS (1985–2002), CINAHL (1982–2002), LISA (1969–2002), PsycInfo (1972–2002), hand search, author contact	All languages	Sampling bias
Hopewell/2002 [18]	8 CS, case study	CMR (2002), MEDLINE (1966–2002), hand search, search reference lists, contact experts	NR	Sampling bias
Lexchin/2003 [22]	30 RC, CS, case study	MEDLINE (1966–2002), EMBASE (1980–2002), CMR, search reference lists, email groups, content experts, author's files, author contact	All languages	Sampling bias
von Elm/2003 [21]	64 RC	MEDLINE, EMBASE, Cochrane Library, CINAHL, BIOSIS, SCI, search reference lists, hand search, author contact, Internet searches[a]	All languages, all formats	Sampling bias
Bekelman/2003 [20]	37 CS, case study	MEDLINE (1980–2002), Web of Science, search reference lists, contact experts, author contact	English, post-1980, published	Sampling bias
[b]Scherer/2005 [23]	79 RC	MEDLINE (up to 2003), EMBASE (up to 2003), CMR (2003), search reference lists, SCI (up to 2003), author's files, word of mouth	NR	Sampling bias
Scherer/1994 [16]	11 RC	MEDLINE, author's files, word of mouth[a]	NR	Sampling bias
Dubben/2005 [24]	26 RC, CS	MEDLINE (1993–2003)	NR	Sampling bias

Abbreviations: RC, retrospective cohort; CS, cross-sectional; CMR, Cochrane Methodology Register; LISA, Library and Information Science Abstracts; DA, Dissertation abstracts; BEI, British Education Index; SIGLE, System for Information on Grey Literature in Europe; NR, not reported; SCI, Science Citation Index; AMED, Allied and Complementary Medicine Database; BIOSIS, Biological abstracts; CINAHL, Cumulative Index to Nursing and Allied Health Literature.

[a] Years of search not reported.

[b] Represents the updated publication of the Scherer 1994 systematic review.

searched electronically [19]. It was concluded that hand searching is valuable in conducting SRs of RCTs, especially for identifying RCTs reported as gray literature, published in languages other than English, and published in journals not indexed in electronic databases.

3.1.1.2.2. Search bias. Hopewell et al. summarized evidence for search bias [19]. The electronic Cochrane Highly Sensitive Search Strategy (HSSS) identified 80% of RCTs compared to 65% for electronic complex searches, and 42% for electronic simple searches.

3.1.1.3. C) Locating studies using reference lists

3.1.1.3.1. Citation bias. One SR examined citation bias [5]. Song et al. summarized nine cross-sectional studies, which found that authors will cite studies that confirm their results and that statistically positive trials are cited more frequently compared to nonpositive trials. For example, one study examining 76 trials found that positive references were cited more frequently (58%) than negative (29%) or neutral (13%) citations [45].

3.1.1.3.2. Multiple/duplicate publication bias. Song et al. reviewed four cross-sectional studies and one retrospective cohort study that examined multiple publication bias [5]. In one study that examined 44 multiple publications of 31 trials, it was determined that the conclusions of multiple publications became more positive over time [46]. In another study, reports with statistically significant results were more likely to generate duplicate publications than those with statistically nonsignificant results [36]. Including duplicate data overestimated the treatment effect in a study that examined a sample of trials [47].

3.1.2. Choosing study biases

This includes *inclusion criteria bias* and *selector bias* [11]. We did not identify a SR that specifically assessed these biases.

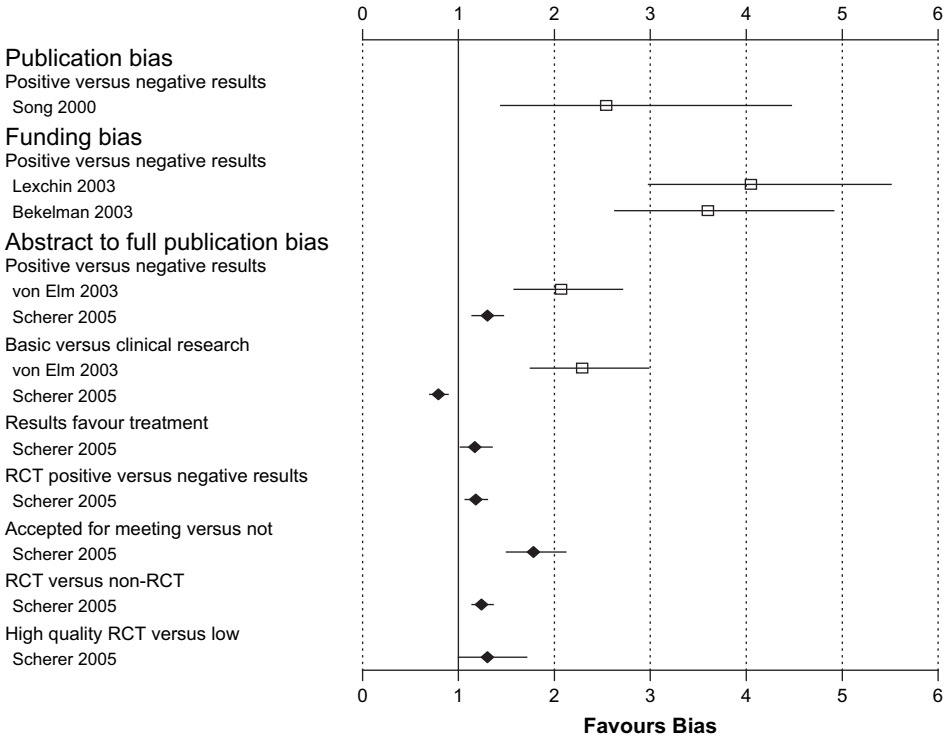

Fig. 2. Forest plot of pooled estimates from random-effects meta-analyses examining bias. Abbreviation: RCT, randomized controlled trial. Legend: Odds ratio □; relative risk ◆.

3.1.3. Obtaining accurate data biases

3.1.3.1. A) Bias by the systematic reviewer. This includes *bias in scoring study quality* and *extractor bias* [11]. We did not identify a SR that specifically assessed these biases.

3.1.3.2. B) Bias due to inaccurate reporting of study results. This includes *study quality bias* and *recording error bias* [11]. We did not identify a SR that specifically assessed these biases.

3.1.3.2.1. Outcome reporting bias. One review examined the phenomenon of outcome reporting bias [5]. The review described two cross-sectional studies and one case study; all found that published trials were more likely to report statistically significant outcomes.

3.1.4. Combining study bias

Bias can occur when statistically combining studies in a SR (e.g., *indirect comparison bias*). We did not identify a SR that specifically assessed this type of bias.

4. Discussion

We identified few SRs documenting the extent of bias that can occur while conducting a SR. Our extensive literature search only identified 10 SRs, one of which was an updated SR. Although few SRs were identified, our results have implications for systematic reviewers. Empirical evidence for publication bias, time-lag bias, abstract to full publication of bias, funding bias, and gray literature bias was identified. These will be described further below.

SRs based only on published material may have exaggerated effect sizes, thus gray literature should be included in SRs. Gray literature, such as conference abstracts, should be sought and included, as evidence suggests that "positive" trials presented as abstracts, oral vs. poster presentations, and RCT designs have a greater likelihood of being published in full. Statistically significant studies tend to be published earlier, overestimating effect sizes of SRs. Therefore, SRs should be routinely updated. Issues surrounding how and when to update SRs were addressed in a recent SR conducted by some of the authors and have been published in the *journal* [48].

Furthermore, at least one database should be searched and although labor intensive, hand searching should be considered whenever feasible. The Cochrane HSSS has the potential to locate a large proportion of RCTs in major English electronic databases and should be used, whenever possible. Although funding bias is omnipresent in the published literature, industry-sponsored publications were of comparable risk of bias to those sponsored by other sources.

We identified SRs examining several biases, yet further investigation into the following is warranted: 1) place of publication bias, 2) country bias, 3) search bias, 4) citation bias, 5) multiple publication bias, and 6) outcome reporting bias (a SR is planned; Dr. P. Williamson, personal

Table 3
Risk of bias in studies included in the systematic reviews

Author/year, reference	Type of risk of bias assessment tool	Risk of bias assessment results
Song/2000 [5]	Not reported	Not reported
Hopewell/2001 [17]	Checklist	Explicit inclusion criteria (2/2), control for clinical differences (1/2) & unclear control (1/2), complete data (1/2), no important flaws (1/2), possible important flaws (1/2)
Hopewell/2002 [19]	Component	Appropriate hand search (17/34) & unclear (17/34), appropriate electronic search (29/34) & unclear (5/34), eligibility hand search agreement (11/34) & unclear (23/34), eligibility electronic search agreement (8/34) & unclear (24/34), comparable search (28/34) & unclear (6/34)
Hopewell/2002 [18]	Checklist	Explicit criteria for gray literature (8/8), gray literature agreement (7/8) & unclear (1/8), data completeness (4/8) & unclear (4/8)
Lexchin/2003 [22]	Not conducted[a]	Not applicable
von Elm/2003 [21]	Formal risk of bias not conducted; instead indicators of methodological quality (e.g., details on follow-up period, databases searched) were assessed	Not explicity reported
Bekelman/2003 [20]	Formal risk of bias not conducted; instead components of the study design of included studies were assessed. For example, the sample size and response rate were assessed for cross-sectional studies.	Not explicity reported
[b]Scherer/2005 [23]	Component	Most studies used an unbiased sample of abstracts, most had at least 2 years of follow-up, and there was fair ascertainment of subsequent publication
Scherer/1994 [16]	Not conducted[a]	Not applicable
Dubben/2005 [24]	Not reported	Not reported

[a] "Not conducted" means the systematic review authors reported that they did not assess the risk of bias in the included studies.
[b] Represents the updated publication of the Scherer 1994 systematic review.

communication). The SRs themselves should be updated, as new evidence may have emerged. Although not a mandate of this review, we believe it is important to explore whether common SR practices do in fact decrease bias, such as having two people independently screen potentially relevant material and scanning the reference lists of the included studies in a SR.

Although many types of bias were covered in the included SRs, gaps in the SR methodological literature were apparent. Our literature search identified additional studies that have yet to be included in a SR of bias. Six cross-sectional studies examining language bias [49–54], 6 retrospective cohort studies examining outcome reporting bias [4,55–59]; 11 cross-sectional studies examining the effects of study risk of bias [52,60–69], although we are aware of a very recent publication examining one component of this [70]; and 6 trials examining whether systematic reviewers should be masked while scoring study quality have not

Table 4
Risk of bias in included systematic reviews (Oxman & Guyatt tool)

Item#	Brief description	Song, 2000 [5]	Hopewell, 2001 [17]	Hopewell, 2002 [18]	Hopewell, 2002 [19]	Lexchin, 2003 [22]	von Elm, 2003 [21]	Bekelman, 2003 [20]	[a]Scherer, 2005 [23]	Scherer, 1994 [16]	Dubben, 2005 [24]
1	Search methods	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
2	Search comprehensiveness	Y	Y	Y	Y	Y	Y	Y	Y	N	N
3	Inclusion criteria	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
4	Bias in study selection	Y	Y	Y	Y	C	C	C	C	C	C
5	Criteria for validity	N	Y	Y	Y	N	P	Y	Y	P	N
6	Appropriate validity items	C	Y	Y	Y	N	C	Y	Y	C	C
7	Combining methods	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
8	Appropriate combining	C	Y	Y	Y	Y	Y	Y	Y	Y	Y
9	Appropriate conclusions	Y	Y	Y	Y	Y	Y	Y	P	Y	C
	[b]Score	4	7	7	7	4	5	6	5	3	2

Abbreviations: Y, yes; N, no; P, partially; C, can't tell.
[a] Represents the major publication of the systematic reviews.
[b] Scoring: Total score is out of 7. A score of 1 means the review has extensive flaws, 2–3 major flaws, 4–5 minor flaws, and 6–7 minimal flaws.

Table 5
Biases examined in the included systematic reviews

Author/year, reference	Details of biases examined	Pooled results of SR: #comparisons, ES: (95% CI)
1. Sampling bias: bias in identifying studies for the systematic review		
A) *Publication bias*		
i) Publication bias		
Song/2000 [5]	Determining the proportion of statistically significant results in the published literature over time	Not conducted
Song/2000 [5]	Examining the association between type of result (e.g., statistically significantly favorable) and publication status (e.g., published)	Four comparisons, adjusted OR for publication bias 2.54 (1.44, 4.47)
Dubben/2005 [24]	Examining the association between type of result and publication status	Not conducted, reported a median reported OR of 2.3
ii) Gray literature bias		
Song/2000 [5]	Examining the association between gray literature and type of result	Not conducted
Hopewell/2002 [18]	Examining the association between gray literature and type of result	Not conducted
iii) Funding bias		
Lexchin/2003 [22]	Examining the association between funding source and publication status	Not conducted
	Examining the association between funding source and type of result	Eighteen comparisons, pharmaceutical sponsorship associated with positive outcomes OR: 4.05 (2.98, 5.51), homogeneity ($P = 0.17$)
	Examining the association between funding source and study quality	Not conducted
Bekelman/2003 [20]	Examining the association between industry sponsorship and type of result	Eight comparisons, industry-sponsored studies associated with positive results OR: 3.60 (2.63, 4.91), homogeneity ($P = 0.75$)
	Examining the association between industry sponsorship and study quality	Not conducted
iv) Time-lag bias		
Song/2000 [5]	Examining the association between time to publication and type of result	Not conducted
Hopewell/2001 [17]	Examining the association between time to publication and type of result	Not conducted
v) Abstract to full publication bias		
Song/2000 [5]	Examining the association between abstract characteristics (e.g., basic science, favorable result) and being published in full	Not conducted
von Elm/2003 [21]	Examining the association between abstract characteristics and being accepted for a conference presentation	46% abstracts submitted to meetings were accepted, acceptance when topic was basic vs. clinical OR: 3.49 (2.50, 4.86) and the outcome was statistically significantly favorable vs. statistically significantly unfavorable OR: 1.67 (1.16, 2.39), heterogeneity NR
	Examining the association between abstract characteristics and being published in full	Abstracts were more likely to be published when topic was basic vs. clinical OR: 2.29 (1.75, 2.98), the outcome was positive vs. negative OR: 2.07 (1.58, 2.71), and it was an oral presentation vs. poster OR: 1.53 (1.15, 2.03), heterogeneity NR
	Examining the association between being rejected for a conference presentation and being published in full	27% abstracts published despite meeting rejection
[a]Scherer/2005 [23]	Examining the association between abstract characteristics and being published in full	44.5% abstracts subsequently published, more likely to be published when there are statistically significant results RR: 1.30 (1.14, 1.47, χ^2: $P = 0.0006$), results favor treatment RR: 1.17 (1.02, 1.35, χ^2: $P = 0.01$), positive results from RCTs RR: 1.18 (1.07, 1.30, χ^2: $P = 0.14$), oral presentation RR: 1.28 (1.09, 1.49, χ^2: $P < 0.0001$), accepted for meeting presentation RR: 1.78 (1.50, 2.12, χ^2: $P < 0.0001$), RCT design OR: 1.24 (1.14, 1.36, χ^2: $P = 0.56$), basic research RR: 0.79 (0.70, 0.89, χ^2: $P = 0.0009$), and higher quality RR: 1.30 (1.00, 1.71, χ^2: $P = 0.68$)

(*Continued*)

Table 5
Continued

Author/year, reference	Details of biases examined	Pooled results of SR: #comparisons, ES: (95% CI)
Scherer/1994 [16]	Examining the association between abstract characteristics and being published in full	Eleven comparisons, 51% abstracts subsequently published, more likely to be published when there are significant results RR: 1.17 (0.99, 1.39) and a sample size above the median RR: 1.48 (1.14, 1.94), homogeneity ($P = 0.01$)
vi) Place of publication bias		
Song/2000 [5]	Examining the association between study characteristics (e.g., topic examined, statistically significantly favorable result) and being published in different journals	Not conducted
vii) Country of conduct bias		
Song/2000 [5]	Examining the association between study characteristics and being conducted by researchers from different countries	Not conducted
viii) Language bias		
Song/2000 [5]	Examining the association between language of publication and study characteristics (e.g., publication status, type of result)	Not conducted
B) *Locating studies using electronic databases*		
i) Indexing bias		
Song/2000 [5]	Examining the association between type of search (e.g., hand searching vs. electronic) and identifying all relevant material	Hand search identified 92–100% of RCTs, whereas electronic searches identified 42–80% of trials
Hopewell/2002 [19]	Examining the association between type of search (e.g., hand searching vs. electronic) and identification of all relevant material	Hand search identified 92–100% of RCTs, whereas electronic searches identified 42–80% of trials
ii) Search bias		
Hopewell/2002 [19]	Examining the association between type of electronic search (e.g., simple vs. complex) and identification of all relevant material	Electronic Cochrane Highly Sensitive Search identified 80% of RCTs, electronic complex searches identified 65% RCTs, and electronic simple searches 42% RCTs, results were judged to be homogeneous but no formal test of heterogeneity conducted
C) *Finding studies using reference lists*		
i) Citation bias		
Song/2000 [5]	Examining the association between searching reference lists and identifying all relevant material	Not conducted
ii) Multiple/duplicate publication bias		
Song/2000 [5]	Examining the association between effect sizes and including duplicate data	Not conducted
2. Choosing study bias: biases in study selection in the systematic review		
A) *Inclusion criteria bias*[b]		
i) Study inclusion bias		
B) *Selector bias*[b]		
3. Obtaining accurate data bias: bias in obtaining accurate data from included studies in the systematic review		
A) *Bias by the systematic reviewer*[b]		
i) Bias in scoring study quality		
ii) Extractor bias		
B) *Bias due to inaccurate reporting of the study results*		
i) Outcome reporting bias		
Song/2000 [5]	Examining the association between outcome characteristic (e.g., statistically significantly favorable) and being reported in the trial.	Not conducted
ii) Study quality bias[b]		
iii) Recording error bias[b]		
4. Combining studies bias: bias that occurs when results are combined		
i) Indirect comparison bias[b]		

Abbreviations: SR, systematic review; MA, meta-analysis; ES, effect size; CI, confidence intervals; RCTs, randomized controlled trials; OR, odds ratio; RR, relative risk.

[a] Represents the updated publication of the Scherer 1994 systematic review.

[b] We did not identify a systematic review that specifically examined this type of bias.

been systematically reviewed [65,71–75]. There may be other gaps in the SR methodological literature that have not yet been fully realized. For example, ''incomplete reporting bias,'' a bias that occurs when studies are omitted from meta-analysis because of incompletely reported information (e.g., measure of dispersion) [76,77]. Even though the literature search identified six SR protocols (Table 1), gaps in the literature are still apparent.

Forty years ago, Archie Cochrane challenged health researchers to systematically review research across all specialties [78]. On the basis of the identified gaps in the SR literature, perhaps methodological reviewers have fallen behind. For example, in the Cochrane Library Issue 2, 2007 there were 4,801 Cochrane reviews, only 19 of which were methodological reviews.

We challenge systematic reviewers to conduct additional high-quality methodological reviews. This will not only inform systematic reviewers in general, but will also impact and inform granting agencies that fund SRs; groups that conduct SRs, such as the Cochrane Collaboration and the Agency for Healthcare Research and Quality Evidence-based Practice Center program; those interested in developing reporting guides for SRs, such as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Group and the Enhancing the Quality and Transparency of Health Research Network (EQUATOR); and texts of how best to conduct SRs.

One of the included SRs had major flaws, four had minor flaws, and four had minimal flaws when assessed for risk of bias. Although 5/9 of the included SRs conducted or reported a meta-analysis, none addressed the potential confounding or effect modification of their results. The effects of a variety of biases on the results should be examined in methodological reviews that conduct a quantitative synthesis, perhaps through stratification, sensitivity analyses, or specifically controlling for confounders and effect modifiers in their analyses. Furthermore, none of the SRs located or included prospective studies as their primary studies. It may be the case that not many methodological prospective studies (e.g., RCTs, prospective cohort studies) examining bias at the SR level have been conducted, which should be a priority research area for methodologists.

Our SR has limitations. Studies published in languages other than English were excluded. Only one independent reviewer screened the titles and abstracts for inclusion. We relied on the study-level risk of bias assessment reported in the SRs. Furthermore, we may have missed pooled analyses that were not based on evidence from a SR (e.g., [79]).

Our findings recommend including unpublished material in SRs, updating SRs periodically, searching more than one database, hand searching for additional material, using the Cochrane HSSS to locate RCT reports, and assessing for publication bias. Further examination of the other types of bias identified in our SR is warranted and the existing empirical evidence should be systematically reviewed.

Acknowledgments

We would like to thank Nick Barrowman for his statistical consultation. The first author would like to thank Drs Moher and Fergusson for supporting this work, as an earlier version was conducted as an assignment for their ''Systematic reviews and meta-analysis'' course at the Department of Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa.

Funding: This research was funded, in part, by Chalmers Research Group. Andrea C. Tricco is funded by a Canadian Institutes of Health Research Canada Graduate Scholarship and by a University of Ottawa National Excellence Scholarship. Dr. Moher is funded by a University of Ottawa Research Chair.

Appendix

Search strategies

MEDLINE (Ovid interface) search strategy

1. bias$.ti,ab.
2. exp ''bias(epidemiology)''/
3. Publication bias/
4. Location bias.mp.
5. Citation bias.mp.
6. Language bias.mp.
7. Reference bias.mp.
8. Multiple publication bias.mp.
9. RANDOMIZED CONTROLLED TRIAL.pt.
10. CONTROLLED CLINICAL TRIAL.pt.
11. RANDOMIZED CONTROLLED TRIALS.sh.
12. RANDOM ALLOCATION.sh.
13. DOUBLE BLIND METHOD.sh.
14. SINGLE BLIND METHOD.sh.
15. or/9-14
16. (ANIMALS not HUMAN).sh.
17. 15 not 16
18. CLINICAL TRIAL.pt.
19. exp CLINICAL TRIALS/
20. (clin$ adj25 trial$).ti,ab.
21. ((singl$ or doubl$ or trebl$ or tripl$) adj25 (blind$ or mask$)).ti,ab.
22. PLACEBO.sh.
23. placebo$.ti,ab.
24. random$.ti,ab.
25. versus.tw.
26. RESEARCH DESIGN.sh.
27. or/18-26
28. 27 not 16
29. 28 not 17
30. 17 or 29
31. (or/1-8) and 30
32. meta-analys$.mp.
33. systematic$ review$.mp.

34. 32 or 33
35. 31 and 34
36. limit 35 to English language

PsycInfo (Ovid interface) search strategy

1. EXPERIMENTER BIAS/ or TEST BIAS/ or CULTURAL TEST BIAS/ or RESPONSE BIAS/
2. bias$.ti,ab.
3. bias$.mp.
4. 2 or 3
5. 4 not 1
6. (meta-analys$ or systematic$ review$).mp.
7. 5 and 6
8. remove duplicates from 7
9. limit 8 to English language

Cochrane Database of Methodological Reviews & Cochrane Methodological Registry (Wiley interface) search strategy

1. Bias* in all fields
2. Meta-analys* or systematic* review* in all fields
3. Meta-analysis or review in publications type
4. MeSH descriptor Meta-analysis explode all tree in MeSH products
5. #1 and (#2 or #3 or #4)

References

[1] Lavis J, Davies H, Oxman A, Denis JL, Golden-Biddle K, Ferlie E. Towards systematic reviews that inform health care management and policy-making. J Health Serv Res Policy 2005;(10 Suppl 1): 35–48.

[2] Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. PLoS Med 2007;4:e78.

[3] Biondi-Zoccai GG, Lotrionte M, Abbate A, Testa L, Remigi E, Burzotta F, et al. Compliance with QUOROM and quality of reporting of overlapping meta-analyses on the role of acetylcysteine in the prevention of contrast associated nephropathy: case study. BMJ 2006;332:202–9.

[4] Silagy CA, Middleton P, Hopewell S. Publishing protocols of systematic reviews. Comparing what was done to what was planned. JAMA 2002;287:2831–4.

[5] Song F, Eastwood AJ, Gilbody S, Duley L, Sutton AJ. Publication and related biases. Health Technol Assess 2000;4:1–115.

[6] Dickersin K, Min YI. Publication bias: the problem that won't go away. Ann N Y Acad Sci 1993;703:135–46.

[7] Dickersin K. How important is publication bias? A synthesis of available data. AIDS Educ Prev 1997;9:15–21.

[8] Cook DJ, Greengold NL, Ellrodt AG, Weingarten SR. The relation between systematic reviews and practice guidelines. Ann Intern Med 1997;127:210–6.

[9] Lavis JN, Posada FB, Haines A, Osei E. Use of research to inform public policymaking. Lancet 2004;364:1615–21.

[10] Lavis JN, Davies HTO, Gruen RL, Walshe K, Farquhar CM. Working within and beyond the Cochrane Collaboration to make systematic reviews more useful to healthcare managers and policymakers. Healthc Policy 2006;1:21–33.

[11] Felson DT. Bias in meta-analytic research. J Clin Epidemiol 1992;45: 885–92.

[12] Peters JL, Sutton AJ, Jones DR, Rushton L, Abrams KR. A systematic review of systematic reviews and meta-analyses of animal experiments with guidelines for reporting. J Environ Sci Health B 2006;41: 1245–58.

[13] Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of reporting of meta-analyses. Lancet 1999;354:1896–900.

[14] Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. J Clin Epidemiol 1991;44:1271–8.

[15] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.

[16] Scherer RW, Dickersin K, Langenberg P. Full publication of results initially presented in abstracts. A meta-analysis. JAMA 1994;272: 158–62.

[17] Hopewell S, Clarke M, Stewart L, Tierney J. Time to publication for results of clinical trials. Cochrane Database Methodol Rev: Rev 2001.

[18] Hopewell S, McDonald S, Clarke M, Egger M. Grey literature in meta-analyses of randomized trials of health care interventions. Cochrane Database Methodol Rev: Rev 2002.

[19] Hopewell S, Clarke M, Lefebvre C, Scherer R. Handsearching versus electronic searching to identify reports of randomized trials. Cochrane Database Methodol Rev: Rev 2002.

[20] Bekelman JE, Li Y, Gross CP. Scope and impact of financial conflicts of interest in biomedical research: a systematic review. JAMA 2003;289:454–65.

[21] von Elm E, Costanza MC, Walder B, Tramer MR. More insight into the fate of biomedical meeting abstracts: a systematic review. BMC Med Res Methodol 2003;3:12.

[22] Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. BMJ 2003;326:1167–70.

[23] Scherer RW, Langenberg P, vonElm E. Full publication of results initially presented in abstracts. Cochrane Database Methodol Rev 2005.

[24] Dubben HH, Beck-Bornholdt HP. Systematic review of publication bias in studies on publication bias. BMJ 2005;331:433–4.

[25] Clarke M, Stewart L, Tierney J, Williamson P. Individual patient data meta-analyses compared with meta-analyses based on aggregate data. Cochrane Database Syst Rev: Protoc 2007.

[26] Leeflang M, McDonald S, Scholten RJPM, Reitsma H, Rutjes AWS. Search strategies to identify diagnostic accuracy studies in MEDLINE and EMBASE. Cochrane Database Syst Rev: Protoc 2007.

[27] McDonald S, Crumley E, Eisinga A, Villanueva E. Search strategies to identify reports of randomized trials in MEDLINE. Cochrane Database Syst Rev: Protoc 2007.

[28] Olsen KL, Hopewell S, Dickersin K, Clarke M, Oxman AD. Publication bias in clinical trials. Cochrane Database Syst Rev: Protoc 2007.

[29] Song F, Altman DG, Glenny A, Eastwood AJ, Deeks JJ. Adjusted indirect comparison for estimating relative effects of competing healthcare interventions. Cochrane Database Syst Rev: Protoc 2007.

[30] Westby M, Clarke M, Hopewell S, Ram F. Masking reviewers at the study inclusion stage in a systematic review of health care interventions. Cochrane Database Syst Rev: Protoc 2007.

[31] Sterling TD. Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa. J Am Stat Assoc 1959;54:30–4.

[32] Sterling TD, Rosenbaum WL, Weinkam JJ. Publication decisions revisited: the effect of the outcome of statistical tests on the decision to publish and vice versa. Am Stat 1995;49:108–12.

[33] Smart RG. The importance of negative results in psychological research. Can Psychol 1964;5:225–32.

[34] Greenwald AG. Consequences of prejudice against the null hypothesis. Psychol Bull 1975;82:1–20.

[35] Moscati R, Jehle D, Ellis D, Fiorello A, Landi M. Positive-outcome bias: comparison of emergency medicine and general medicine literatures. Acad Emerg Med 1994;1:267−71.
[36] Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. Lancet 1991;337:867−72.
[37] Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. JAMA 1992;267:374−8.
[38] Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. BMJ 1997;315:640−5.
[39] Ioannidis JP. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. JAMA 1998;279:281−6.
[40] Rothwell PM, Robertson G. Meta-analyses of randomised controlled trials. Lancet 1997;350:1181−2.
[41] Ben-Shlomo Y, Smith GD. "Place of publication" bias? BMJ 1994;309:274.
[42] Simes RJ. Confronting publication bias: a cohort design for meta-analysis. Stat Med 1987;6:11−29.
[43] Ottenbacher K, DiFabio RP. Efficacy of spinal manipulation/mobilization therapy. A meta-analysis. Spine 1985;10:833−7.
[44] Vickers A, Goyal N, Harland R, Rees R. Do certain countries produce only positive results? A systematic review of controlled trials. Control Clin Trials 1998;19:159−66.
[45] Gotzsche PC. Reference bias in reports of drug trials. Br Med J (Clin Res Ed) 1987;295:654−6.
[46] Gotzsche PC. Multiple publication of reports of drug trials. Eur J Clin Pharmacol 1989;36:429−32.
[47] Tramer MR, Reynolds DJ, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. BMJ 1997;315:635−40.
[48] Moher D, Tsertsvadze A, Tricco AC, Eccles M, Grimshaw J, Sampson M, et al. Systematic review identified few methods and strategies describing when and how to update systematic reviews. J Clin Epidemiol 2007;60:1095−104.
[49] Moher D, Fortin P, Jadad AR, Juni P, Klassen T, Le LJ, et al. Completeness of reporting of trials published in languages other than English: implications for conduct and reporting of systematic reviews. Lancet 1996;347:363−6.
[50] Egger M, Zellweger-Zahner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. Lancet 1997;350:326−9.
[51] Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. Health Technol Assess 2003;7:1−90.
[52] Fergusson D, Laupacis A, Salmi LR, McAlister FA, Huet C. What should be included in meta-analyses? An exploration of methodological issues using the ISPOT meta-analyses. Int J Technol Assess Health Care 2000;16:1109−19.
[53] Juni P, Holenstein F, Sterne J, Bartlett C, Egger M. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. Int J Epidemiol 2002;31:115−23.
[54] Pham B, Klassen TP, Lawson ML, Moher D. Language of publication restrictions in systematic reviews gave different results depending on whether the intervention was conventional or complementary. J Clin Epidemiol 2005;58:769−76.
[55] Hahn S, Williamson PR, Hutton JL, Garner P, Flynn EV. Assessing the potential for bias in meta-analysis due to selective reporting of subgroup analyses within studies. Stat Med 2000;19:3325−36.
[56] Chan AW, Krleza-Jeric K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. CMAJ 2004;171:735−40.
[57] 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.
[58] Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. BMJ 2005;330:753.
[59] Williamson PR, Gamble C, Altman DG, Hutton JL. Outcome selection bias in meta-analysis. Stat Methods Med Res 2005;14:515−24.
[60] Chalmers TC, Celano P, Sacks HS, Smith H Jr. Bias in treatment assignment in controlled clinical trials. N Engl J Med 1983;309:1358−61.
[61] Miller JN, Colditz GA, Mosteller F. How study design affects outcomes in comparisons of therapy. II: Surgical. Stat Med 1989;8:455−66.
[62] Colditz GA, Miller JN, Mosteller F. How study design affects outcomes in comparisons of therapy. I: Medical. Stat Med 1989;8:441−54.
[63] Emerson JD, Burdick E, Hoaglin DC, Mosteller F, Chalmers TC. An empirical study of the possible relation of treatment differences to quality scores in controlled randomized clinical trials. Control Clin Trials 1990;11:339−52.
[64] 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.
[65] Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet 1998;352:609−13.
[66] Egger M, Juni P, Bartlett C, Holenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. Health Technol Assess 2003;7:1−76.
[67] Khan KS, Daya S, Jadad A. The importance of quality of primary studies in producing unbiased systematic reviews. Arch Intern Med 1996;156:661−6.
[68] Balk EM, Bonis PA, Moskowitz H, Schmid CH, Ioannidis JP, Wang C, et al. Correlation of quality measures with estimates of treatment effect in meta-analyses of randomized controlled trials. JAMA 2002;287:2973−82.
[69] van Nieuwenhoven CA, Buskens E, van Tiel FH, Bonten MJ. Relationship between methodological trial quality and the effects of selective digestive decontamination on pneumonia and mortality in critically ill patients. JAMA 2001;286:335−40.
[70] Pildal J, Hrobjartsson A, Jorgensen K, Hilden J, Altman D, Gotzsche P. Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. Int J Epidemiol 2007.
[71] 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.
[72] Verhagen AP, de Vet HC, Vermeer F, Widdershoven JW, de Bie RA, Kessels AG, et al. The influence of methodologic quality on the conclusion of a landmark meta-analysis on thrombolytic therapy. Int J Technol Assess Health Care 2002;18:11−23.
[73] Berlin JA. Does blinding of readers affect the results of meta-analyses? University of Pennsylvania Meta-analysis Blinding Study Group. Lancet 1997;350:185−6.
[74] Clark HD, Wells GA, Huet C, McAlister FA, Salmi LR, Fergusson D, et al. Assessing the quality of randomized trials: reliability of the Jadad scale. Control Clin Trials 1999;20:448−52.
[75] Berard A, Andreu N, Tetrault J, Niyonsenga T, Myhal D. Reliability of Chalmers' scale to assess quality in meta-analyses on pharmacological treatments for osteoporosis. Ann Epidemiol 2000;10:498−503.
[76] Caldwell DJ, Armstrong TW, Barone NJ, Suder JA, Evans MJ. Lessons learned while compiling a quantitative exposure database from the published literature. Appl Occup Environ Hyg 2001;16:174−7.
[77] Gamble C, Hollis S. Uncertainty method improved on best-worst case analysis in a binary meta-analysis. J Clin Epidemiol 2005;58:579−88.

[78] Cochrane AL. 1931–1971: a critical review, with particular reference to the medical profession. Medicines for the year 2000. London: Office of Health Economics; 1979; 1–11.
[79] Gluud LL. Bias in clinical intervention research. Am J Epidemiol 2006;163:493–501.
[80] Jadad AR, Rennie D. The randomized controlled trial gets a middle-aged checkup. JAMA 1998;279:319–20.
[81] Gregoire G, Derderian F, Le LJ. Selecting the language of the publications included in a meta-analysis: is there a Tower of Babel bias? J Clin Epidemiol 1995;48:159–63.

Glossary of biases

Publication bias:
Occurs when investigators, reviewers, and editors submit or accept manuscripts for publication based on the direction or strength of the study findings [7].
Gray literature bias:
Occurs when the results reported in journal articles are systematically different from those presented in other reports, such as working papers, dissertations, or conference abstracts [5].
Funding bias:
Biases in the design, outcome, and reporting of industry-sponsored research to show that a drug shows a favorable outcome [22].
Time-lag bias:
Occurs when the speed of publication depends on the direction and strength of the trial results [80].
Abstract to full publication bias:
Occurs when the full publication of studies that have been initially presented at conferences or in other informal formats is dependent on the direction and/or strength of their findings [5].
Place of publication bias:
Occurs when a journal is more enthusiastic toward publishing articles about a given hypothesis than other journals because of editorial policy or readers' preference [41].
Country of conduct bias:
Occurs when the country of publication is associated with the strength or direction of research findings [5].
Language bias:
Occurs when languages of publication depend on the direction and strength of the study results [81].
Indexing bias:
Occurs when there is biased indexing of published studies in literature databases [11].
Search bias:
Occurs when there is a bias in captured studies resulting from an inadequate or incomplete search [11].
Citation bias:
Occurs when the chance of a study being cited by others is associated with its results. Therefore, retrieving literature from scanning reference lists may produce a biased sample of articles, rendering the conclusions of an article less reliable [5], [45].
Multiple/duplicate publication bias:
Occurs when studies with significant or supportive results are more likely to generate multiple publications than studies with nonsignificant or unsupportive results. It can be classified as overt or covert, the latter being more difficult to deal with in systematic reviews (SRs) [5], [47].
Inclusion criteria bias:
Occurs when the inclusion criteria of a review purposely exclude important studies that the reviewer knows of [11].
Selector bias:
Occurs when the inclusion criteria are not specific enough, leaving the reviewer free to choose studies, which may be susceptible to bias [11].
Bias in scoring study quality:
Occurs when the systematic reviewer systematically scores studies published by their peers or in high-impact journals as being more methodologically rigorou [11].
Extractor bias:
Occurs when the data are not extracted accurately from the study [11].
Study quality bias:
Occurs when studies of lower or higher quality are associated with positive or favorable results [11].
Recording error bias:
Occurs when the actual study results and the recorded results in the published paper differ [11].
Outcome reporting bias:
Occurs when a study in which multiple outcomes were measured reports only those that are significant [5].
Indirect comparison bias:
Occurs when indirect comparisons rather than direct comparisons are used to combine results in a SR.