



Cochrane
Library

Cochrane Database of Systematic Reviews

Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions (Review)

Page MJ, McKenzie JE, Kirkham J, Dwan K, Kramer S, Green S, Forbes A

Page MJ, McKenzie JE, Kirkham J, Dwan K, Kramer S, Green S, Forbes A.

Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions.

Cochrane Database of Systematic Reviews 2014, Issue 10. Art. No.: MR000035.

DOI: 10.1002/14651858.MR000035.pub2.

www.cochranelibrary.com

TABLE OF CONTENTS

| | |
|---|----|
| HEADER | 1 |
| ABSTRACT | 1 |
| PLAIN LANGUAGE SUMMARY | 2 |
| BACKGROUND | 3 |
| OBJECTIVES | 4 |
| METHODS | 4 |
| RESULTS | 8 |
| Figure 1. | 9 |
| Figure 2. | 11 |
| Figure 3. | 12 |
| Figure 4. | 13 |
| Figure 5. | 14 |
| Figure 6. | 17 |
| DISCUSSION | 17 |
| AUTHORS' CONCLUSIONS | 20 |
| ACKNOWLEDGEMENTS | 20 |
| REFERENCES | 20 |
| CHARACTERISTICS OF STUDIES | 26 |
| DATA AND ANALYSES | 42 |
| Analysis 1.1. Comparison 1 Association between outcome discrepancies and statistical significance, Outcome 1 Outcome "added/upgraded". | 42 |
| Analysis 1.2. Comparison 1 Association between outcome discrepancies and statistical significance, Outcome 2 Outcome "downgraded". | 43 |
| APPENDICES | 43 |
| FEEDBACK | 49 |
| WHAT'S NEW | 49 |
| CONTRIBUTIONS OF AUTHORS | 50 |
| DECLARATIONS OF INTEREST | 50 |
| SOURCES OF SUPPORT | 50 |
| DIFFERENCES BETWEEN PROTOCOL AND REVIEW | 50 |
| INDEX TERMS | 51 |

[Methodology Review]

Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions

Matthew J Page¹, Joanne E McKenzie¹, Jamie Kirkham², Kerry Dwan², Sharon Kramer¹, Sally Green¹, Andrew Forbes³

¹School of Public Health & Preventive Medicine, Monash University, Melbourne, Australia. ²Department of Biostatistics, University of Liverpool, Liverpool, UK. ³Department of Epidemiology & Preventive Medicine, Monash University, Melbourne, Australia

Contact address: Matthew J Page, School of Public Health & Preventive Medicine, Monash University, The Alfred Centre, 99 Commercial Road, Melbourne, Victoria, 3004, Australia. matthew.page@monash.edu.

Editorial group: Cochrane Methodology Review Group.

Publication status and date: New, published in Issue 10, 2014.

Review content assessed as up-to-date: 8 May 2013.

Citation: Page MJ, McKenzie JE, Kirkham J, Dwan K, Kramer S, Green S, Forbes A. Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions. *Cochrane Database of Systematic Reviews* 2014, Issue 10. Art. No.: MR000035. DOI: 10.1002/14651858.MR000035.pub2.

Copyright © 2014 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Systematic reviews may be compromised by selective inclusion and reporting of outcomes and analyses. Selective inclusion occurs when there are multiple effect estimates in a trial report that could be included in a particular meta-analysis (e.g. from multiple measurement scales and time points) and the choice of effect estimate to include in the meta-analysis is based on the results (e.g. statistical significance, magnitude or direction of effect). Selective reporting occurs when the reporting of a subset of outcomes and analyses in the systematic review is based on the results (e.g. a protocol-defined outcome is omitted from the published systematic review).

Objectives

To summarise the characteristics and synthesise the results of empirical studies that have investigated the prevalence of selective inclusion or reporting in systematic reviews of randomised controlled trials (RCTs), investigated the factors (e.g. statistical significance or direction of effect) associated with the prevalence and quantified the bias.

Search methods

We searched the Cochrane Methodology Register (to July 2012), Ovid MEDLINE, Ovid EMBASE, Ovid PsycINFO and ISI Web of Science (each up to May 2013), and the US Agency for Healthcare Research and Quality (AHRQ) Effective Healthcare Program's Scientific Resource Center (SRC) Methods Library (to June 2013). We also searched the abstract books of the 2011 and 2012 Cochrane Colloquia and the article alerts for methodological work in research synthesis published from 2009 to 2011 and compiled in *Research Synthesis Methods*.

Selection criteria

We included both published and unpublished empirical studies that investigated the prevalence and factors associated with selective inclusion or reporting, or both, in systematic reviews of RCTs of healthcare interventions. We included empirical studies assessing any type of selective inclusion or reporting, such as investigations of how frequently RCT outcome data is selectively included in systematic reviews based on the results, outcomes and analyses are discrepant between protocol and published review or non-significant outcomes are partially reported in the full text or summary within systematic reviews.

Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions (Review)

Copyright © 2014 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Data collection and analysis

Two review authors independently selected empirical studies for inclusion, extracted the data and performed a risk of bias assessment. A third review author resolved any disagreements about inclusion or exclusion of empirical studies, data extraction and risk of bias. We contacted authors of included studies for additional unpublished data. Primary outcomes included overall prevalence of selective inclusion or reporting, association between selective inclusion or reporting and the statistical significance of the effect estimate, and association between selective inclusion or reporting and the direction of the effect estimate. We combined prevalence estimates and risk ratios (RRs) using a random-effects meta-analysis model.

Main results

Seven studies met the inclusion criteria. No studies had investigated selective inclusion of results in systematic reviews, or discrepancies in outcomes and analyses between systematic review registry entries and published systematic reviews. Based on a meta-analysis of four studies (including 485 Cochrane Reviews), 38% (95% confidence interval (CI) 23% to 54%) of systematic reviews added, omitted, upgraded or downgraded at least one outcome between the protocol and published systematic review. The association between statistical significance and discrepant outcome reporting between protocol and published systematic review was uncertain. The meta-analytic estimate suggested an increased risk of adding or upgrading (i.e. changing a secondary outcome to primary) when the outcome was statistically significant, although the 95% CI included no association and a decreased risk as plausible estimates (RR 1.43, 95% CI 0.71 to 2.85; two studies, $n = 552$ meta-analyses). Also, the meta-analytic estimate suggested an increased risk of downgrading (i.e. changing a primary outcome to secondary) when the outcome was statistically significant, although the 95% CI included no association and a decreased risk as plausible estimates (RR 1.26, 95% CI 0.60 to 2.62; two studies, $n = 484$ meta-analyses). None of the included studies had investigated whether the association between statistical significance and adding, upgrading or downgrading of outcomes was modified by the type of comparison, direction of effect or type of outcome; or whether there is an association between direction of the effect estimate and discrepant outcome reporting.

Several secondary outcomes were reported in the included studies. Two studies found that reasons for discrepant outcome reporting were infrequently reported in published systematic reviews (6% in one study and 22% in the other). One study (including 62 Cochrane Reviews) found that 32% (95% CI 21% to 45%) of systematic reviews did not report all primary outcomes in the abstract. Another study (including 64 Cochrane and 118 non-Cochrane reviews) found that statistically significant primary outcomes were more likely to be completely reported in the systematic review abstract than non-significant primary outcomes (RR 2.66, 95% CI 1.81 to 3.90). None of the studies included systematic reviews published after 2009 when reporting standards for systematic reviews (Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) Statement, and Methodological Expectations of Cochrane Intervention Reviews (MECIR)) were disseminated, so the results might not be generalisable to more recent systematic reviews.

Authors' conclusions

Discrepant outcome reporting between the protocol and published systematic review is fairly common, although the association between statistical significance and discrepant outcome reporting is uncertain. Complete reporting of outcomes in systematic review abstracts is associated with statistical significance of the results for those outcomes. Systematic review outcomes and analysis plans should be specified prior to seeing the results of included studies to minimise post-hoc decisions that may be based on the observed results. Modifications that occur once the review has commenced, along with their justification, should be clearly reported. Effect estimates and CIs should be reported for all systematic review outcomes regardless of the results. The lack of research on selective inclusion of results in systematic reviews needs to be addressed and studies that avoid the methodological weaknesses of existing research are also needed.

PLAIN LANGUAGE SUMMARY

Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions

A systematic review summarises evidence from multiple studies to answer a specific research question (e.g. what are the benefits and harms of a particular intervention for a particular health condition?). Often, there are many outcomes that systematic review authors could report to address their research question (e.g. pain, disability and quality of life for patients with musculoskeletal conditions) and many different results available for a particular outcome (e.g. a study might measure pain using three different scales at four time points). If the decision about which outcomes to investigate in a systematic review is made based on the results for those outcomes in

the eligible studies, this may lead to bias. While, if the decision about which outcomes to report in a systematic review and the ways to report them is based on the results, this may mislead users of the systematic review.

This methodology review summarises the findings of studies examining the inclusion of results and reporting of outcomes in systematic reviews. We searched for studies indexed in electronic bibliographic databases up to May 2013. We included seven studies and found that outcomes investigated and reported in systematic reviews were often changed between the protocol and published systematic review. We also found that it was unclear whether the decision to make these changes was related to how statistically convincing the treatment effect for that outcome was. More studies are needed to confirm if this relationship exists. Also, one study found that some systematic reviews did not report all of the most important outcomes in the abstract of the review. Another study found that outcomes with a more statistically convincing result were more likely to be completely reported in the abstract than other outcomes. The studies that we included were limited to systematic reviews published before 2009. New studies are needed to examine the inclusion of results and reporting of outcomes in more recent systematic reviews.

BACKGROUND

Systematic reviews and meta-analyses of intervention studies provide evidence to decision makers on the benefits and harms of healthcare interventions, keep clinicians up-to-date, identify gaps in knowledge and provide recommendations for future research (Higgins 2011b; Moher 2009). However, various practices may affect the validity of systematic reviews. These practices include selective publication of studies (Begg 1988; Dickersin 1990) and selective reporting of a subset of outcomes and analyses within studies based on the results (e.g. statistical significance, magnitude or direction of effect). We refer to the latter as selective reporting (Hutton 2000; Song 2010).

Selective reporting can occur in various ways in randomised controlled trials (RCTs). Examples include omission of all data for a measured outcome, reporting data for only a subset of the time points measured, and reporting only subgroup analyses or subscale data. Additional problems include the failure to report measures of variation (e.g. standard errors (SEs) or 95% confidence intervals (CIs)) or exact P values for non-significant outcomes, and modification of 'primary' and 'secondary' outcome labels between registry entries, protocols and publications (Chan 2004a; Chan 2004b; Chan 2005; Dwan 2008; Dwan 2014; Mathieu 2009). Certain types of selective reporting in RCTs can lead to bias in the results of meta-analyses (Kirkham 2010a; Williamson 2005a; Williamson 2005b).

Empirical evidence of selective reporting in RCTs exists. In two landmark studies, Chan and colleagues compared outcomes reported in RCT protocols to the methods and results sections of publications. They found that 71% of RCTs in one study and 88% in another had at least one unreported efficacy outcome (Chan 2004a; Chan 2004b). A systematic review of nine cohorts of RCTs found that at least one primary outcome was added, omitted or changed (upgraded from secondary to primary or downgraded

from primary to secondary) in 4% to 50% of publications when compared with the registry entry or protocol. In addition, statistically significant outcomes were more often completely reported compared with non-significant outcomes (range of odds ratios 2.4 to 4.7) where completely reported was defined as reporting sufficient data for inclusion in a meta-analysis (Dwan 2011).

Description of the problem or issue

Most research on selective reporting has been undertaken at the RCT level (Norris 2013; Page 2013a). However, there is potential for similar processes to occur at the systematic review level, particularly when multiplicity of outcome data exists in RCT reports. Examples of multiplicity of data include multiple measurement scales, multiple time points and multiple analyses (e.g. when both final and change from baseline values, or both intention-to-treat and per-protocol analyses are reported for a particular outcome) (Bender 2008; Tendal 2011). When multiplicity of data is present, systematic review authors may include a subset of outcome data in the review. There can be reasonable justifications for including a subset of outcome data in systematic reviews and meta-analyses (e.g. choosing to include data collected using validated measurement scales only, or excluding surrogate outcomes that are of limited clinical relevance), but if the choice is made based on the results (which we refer to as 'selective inclusion of results') the corresponding meta-analytic effect estimate may be biased (Page 2013a; Tendal 2009; Tendal 2011).

After inclusion in a systematic review, outcomes and analyses may be selectively reported in the same ways as occurs in RCTs (Kirkham 2010b; Moher 2007). For example, systematic review authors may not report results of all primary outcomes in the abstract of the review, or may modify the description of outcomes from primary to secondary or vice versa. Selective reporting can

misrepresent the available evidence and mislead the users of systematic reviews regarding the importance of particular outcomes (Chalmers 1990; Liberati 2009; Moher 2009).

There are important differences between empirical investigations of selective reporting in systematic reviews and selective reporting in RCTs. Unlike RCTs, which involve prospective recruitment of participants and collection of outcome data, most systematic reviews are retrospective by nature, in that the studies included are usually identified after they have been completed and their results reported. This means that the outcomes and analyses reported in systematic reviews are dependent on the data available in the included studies. Therefore, when investigating bias due to selective inclusion or reporting in systematic reviews, the extent to which changes to planned outcomes and analyses occurred because none of the included studies measured the necessary outcome data (rather than because of the nature of the results) requires consideration by investigators.

Why it is important to do this review

Several empirical studies have assessed the prevalence of selective inclusion and reporting in systematic reviews of RCTs, and the factors associated with the prevalence, but we are unaware of any systematic reviews of such studies. If selective inclusion and reporting are found to occur frequently, interventions may be necessary to improve the conduct and reporting of systematic reviews. Examples of such interventions include: increasing review authors' awareness of guidelines for the reporting of outcomes in systematic reviews, particularly the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement (Liberati 2009; Moher 2009) and the Methodological Expectations of Cochrane Intervention Reviews (MECIR) reporting standards (Chandler 2012); more detailed advice in guidance documents such as the *Cochrane Handbook for Systematic Reviews of Interventions* regarding these potential sources of bias (Higgins 2011a); the development of core outcome sets for a range of clinical conditions (Clarke 2007; COMET Initiative; Kirkham 2013); and registration of detailed protocols for all systematic reviews of healthcare interventions (Booth 2010; PLoS 2011).

OBJECTIVES

To summarise the characteristics and synthesise the results of empirical studies that have investigated the prevalence of selective inclusion or reporting in systematic reviews of RCTs, investigated the factors (e.g. statistical significance or direction of effect) associated with the prevalence and quantified the bias.

METHODS

Criteria for considering studies for this review

Types of studies

We included both published and unpublished empirical studies that investigated the prevalence and factors associated with selective inclusion or reporting, or both, in systematic reviews of RCTs of healthcare interventions. If empirical studies included a mixture of systematic reviews of RCTs and non-randomised studies (NRS) and reported data separately, we only included the findings for systematic reviews of RCTs. If it was not possible to separate the data, we contacted the study authors. If the data could not be separated, we presented the results but did not include them in the quantitative syntheses.

We included empirical studies that comprised a sample or a complete set of systematic reviews (e.g. all systematic reviews registered or published during a specific time period). The empirical studies could have assessed any type of selective inclusion or reporting, such as investigations of how frequently: (1) RCT outcome data is selectively included in systematic reviews based on the results; (2) outcomes and analyses are discrepant between protocol and published review; or (3) non-significant outcomes are partially reported in the full text or summary within systematic reviews. We defined the full text of the systematic review as the text reported in the results section, tables, forest plots and data available via online appendices. Summaries within systematic reviews included the abstract, plain language summary or 'Summary of findings' tables. We excluded empirical studies assessing discrepancies in information other than outcomes and analyses, e.g. search strategy or inclusion criteria, as discrepancies of this nature were beyond the scope of our review.

Prior to the launch in February 2011 of PROSPERO, an international online prospective register of systematic reviews hosted by the Centre for Reviews and Dissemination (University of York, United Kingdom), access to systematic review protocols was generally limited to organisations such as The Cochrane Collaboration, the Campbell Collaboration and the Joanna Briggs Institute (Booth 2010; Booth 2011). Therefore, we anticipated that it was unlikely that empirical studies that included systematic reviews other than those coming from these organisations would exist at the time of our search. We also anticipated that empirical studies comparing systematic review registry entries to published systematic reviews would be unlikely to exist at this time. However, such studies may be included in future updates of this Cochrane Review. While these issues may limit the generalisability of some of the empirical studies identified, these issues are unlikely to apply to empirical studies investigating selective inclusion of RCT outcome data in systematic reviews or selective partial reporting in systematic reviews, because both of these types of selective inclusion and reporting can be assessed in systematic reviews without a systematic review protocol or registry entry.

Types of data

We included estimates of the prevalence of selective inclusion or reporting in systematic reviews of RCTs and estimates of the association between selective inclusion or reporting and any factors predictive of this (e.g. statistical significance or direction of the effect) as assessed by the investigators who did the original empirical studies. We anticipated that these factors may have been defined by investigators differently, but included any empirical study regardless of the definitions used. The terms “full” or “complete” versus “partial” reporting in either the full text or summary of the systematic review were also anticipated to be defined variously by investigators, but we included any empirical study regardless of the definitions used.

Types of methods

We focused on five types of selective inclusion and reporting of outcomes and analyses in this review. These included:

- selective inclusion of RCT outcome data in systematic reviews;
- discrepancies between systematic review registry entries and published systematic reviews;
- discrepancies between systematic review protocols and published systematic reviews;
- discrepancies between the full text and the summaries (i.e. abstract, plain language summary or ‘Summary of findings’ table) in systematic reviews;
- selective partial reporting in systematic reviews.

Types of outcome measures

For each of the five types of methods, we included three primary outcomes:

1. Overall prevalence of selective inclusion or reporting;
2. Whether the overall prevalence of selective inclusion or reporting is associated with statistical significance of the effect estimate;
3. Whether the overall prevalence of selective inclusion or reporting is associated with direction of the effect estimate.

We also included four secondary outcomes:

1. Prevalence of specific examples of selective inclusion or reporting (see [Appendix 1](#) for examples that we anticipated might be reported in empirical studies);
2. Whether the prevalence of each specific example of selective inclusion or reporting is associated with statistical significance of the effect estimate;
3. Whether the prevalence of each specific example of selective inclusion or reporting is associated with the direction of the effect estimate;
4. Whether the overall prevalence or prevalence of specific examples of selective inclusion or reporting is associated with other factors investigated by the empirical study investigators (e.g. clinical area, funding of the systematic review).

We did not exclude studies based on the outcomes reported.

Search methods for identification of studies

Electronic searches

We searched the following electronic databases:

- Cochrane Methodology Register (to July 2012, which was the last month in which new records were added to this database);
- Ovid MEDLINE (January 1946 to May 2013);
- Ovid EMBASE (January 1980 to May 2013);
- Ovid PsycINFO (January 1806 to May 2013);
- ISI Web of Science (January 1898 to May 2013).

The search strategies for each database are reported in [Appendix 2](#). We also searched the US Agency for Healthcare Research and Quality (AHRQ) Effective Healthcare Program’s Scientific Resource Center (SRC) Methods Library ([SRC Methods Library 2013](#)) on 5th June 2013, using the following Search Descriptors: “Bias - outcome selection”, “Bias - Publication”, “Bias - Reporting”, “Reporting”, “Registries - Systematic Reviews” and “Synthesis, Quantitative - Reporting”.

Searching other resources

We searched the abstract books of the 2011 and 2012 Cochrane Colloquia (which were not indexed in any electronic databases at the time of the search). We also searched the article alerts for methodological work in research synthesis published from 2009 to 2011, which are available in *Research Synthesis Methods* ([Hafdahl 2010b](#); [Hafdahl 2010a](#); [Hafdahl 2011a](#); [Hafdahl 2011b](#); [Hafdahl 2012](#)). One review author (MJP) screened the reference lists of all included studies and any relevant reviews identified from the search. One review author (MJP) contacted the authors of all included studies and relevant reviews, as well as individuals with content expertise, to assist with the identification of published, unpublished and ongoing studies.

Data collection and analysis

Selection of studies

Two of four review authors (MJP and either JK, KD or SK) independently screened the titles and available abstracts of all studies identified by the search against the inclusion criteria (see [Criteria for considering studies for this review](#)), and excluded any clearly irrelevant studies. We did not exclude studies based on the language of the publication. Two review authors (MJP and SK) independently assessed full-text copies of reports of potentially eligible studies. The authors resolved disagreements regarding study

inclusion by discussion, involving a third review author (JEM) when required.

Data extraction and management

Two of four review authors (MJP and either JK, KD or SK) independently extracted data using a data extraction form developed for this Cochrane Review. Any discrepancies between the authors were resolved through discussion until consensus was reached or by arbitration of a third review author (JEM) when required. All the data extractors pilot tested the data extraction form and modified it accordingly before use. One review author (MJP) compiled all comparisons and entered outcome data into Review Manager 5 (RevMan 2011). One review author (JEM) cross-checked the data entered. For studies where relevant outcome data were not reported, one review author (MJP) requested further information from the study investigators. When unsuccessful, we included the study in the review and fully described it (e.g. using tables), but did not include it in any quantitative syntheses.

We extracted the following data from each study:

- characteristics of the study, in particular the types of selective inclusion or reporting investigated, the number of included systematic reviews, methods for selecting the systematic reviews, areas of health care addressed, range of years of publication of the systematic reviews, proportion of systematic reviews that were Cochrane Reviews and methodological quality of the systematic reviews (however assessed by the study authors);
- data on estimates of prevalence and association, as specified under [Types of outcome measures](#);
- the definition of factors investigated for their association with selective inclusion or reporting (e.g. statistical significance defined as $P < 0.05$);
- the definition of “full” or “partial” reporting used by study investigators who assessed this practice;
- any confounding variables assessed by the study authors (e.g. funding type).

Assessment of risk of bias in included studies

There is no standard tool available to evaluate the risk of bias of empirical studies eligible for inclusion in this review. We used the following criteria:

1. What is the risk of selection bias in the empirical study? Low risk of bias: the empirical study included all systematic reviews registered during a specified time period (where registration is defined as registration of the review in an online database or publication of a protocol for the review), or included a random sample of systematic reviews. High risk of bias: the empirical study included a non-random sample of systematic reviews. Unclear risk of bias: the sampling frame for the sample of systematic reviews of RCTs is unclear.
2. What is the risk of selective reporting bias in the empirical study? Low risk of bias: all comparisons and outcomes reported

in the protocol for the empirical study are fully reported in the results section of the publication. High risk of bias: not all comparisons and outcomes reported in the protocol for the empirical study are fully reported in the results section of the publication. Unclear risk of bias: it is unclear if all comparisons and outcomes are fully reported in the results section of the publication (e.g. because a protocol for the empirical study is not available) (Dwan 2011).

Two of four authors (MJP and either JK, KD or SK) rated each criterion independently. Any discrepancies between the authors were resolved through discussion until consensus was reached or by arbitration of a third author (JEM) when required.

Measures of the effect of the methods

The measures of prevalence and associations between a factor and selective inclusion or reporting were dependent on the summary statistics reported by the empirical study investigators, and we reported the data as available. For empirical studies that reported estimates of prevalence, we reported percentages with 95% CIs. We have presented possible prevalence estimates that we considered including in [Appendix 1](#). For empirical studies investigating the association between a factor and selective inclusion or reporting, we reported risk ratios (RRs) with 95% CIs. We standardised outcomes so that they have a consistent meaning. For example, RRs > 1 denote a higher risk of selective inclusion or reporting bias. If the empirical study investigators reported odds ratios (ORs), we converted these to RRs using the formula provided in section 12.5.4.4 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Schünemann 2011). For studies that quantified bias in meta-analytic estimates due to selective inclusion, we reported the estimates of bias as reported in the empirical studies.

Unit of analysis issues

Within the systematic reviews evaluated in an empirical study, there is potential for overlap of included RCTs. We contacted the authors of the empirical studies to enquire how they dealt with this issue of RCT overlap. We reported whether the issue was addressed and discussed its likely impact on the results of the empirical study. Furthermore, we anticipated that there might be some overlap in the systematic reviews included in the empirical studies in our review. Based on information regarding the types of reviews (i.e. Cochrane versus non-Cochrane), years of publication, clinical condition and types of outcomes (e.g. dichotomous versus continuous), we reported how likely the potential for overlap is and how it may have impacted on the results of our review. If we suspected that two studies were likely to have a substantial proportion of overlapping systematic reviews, we performed sensitivity analyses by removing the study with the smaller number of included systematic reviews from the meta-analysis.

Dealing with missing data

If the empirical studies had missing information on the characteristics of the included studies (e.g. methods for selecting the systematic reviews) or missing outcome data (e.g. measure of variation for a RR denoting the association between statistical significance and discrepant outcome reporting), we contacted the authors. We did not plan to impute any missing outcome data.

Assessment of heterogeneity

We assessed methodological heterogeneity by determining whether the characteristics of the included empirical studies were similar, particularly in terms of the areas of health care addressed by the systematic reviews and type of methodological comparisons undertaken. We assessed statistical heterogeneity by inspecting the forest plots and by calculating the I^2 statistic with 95% CIs (Deeks 2011; Higgins 2002). We calculated 95% CIs for the I^2 statistic using the non-central χ^2 approximation implemented in the Stata module *heterogi* (this module can only be used when there are more than two studies included in a meta-analysis) (Orsini 2006).

Assessment of reporting biases

To assess small study effects, we planned to generate funnel plots if at least 10 empirical studies examining the same methodological comparison and outcome met our inclusion criteria (Sterne 2011). To assess selective reporting in the included empirical studies, we compared outcomes reported in the empirical study protocol (if available) with outcomes reported in the results section of the publication. We searched for empirical study protocols in the electronic databases listed above (see [Electronic searches](#)). If a published protocol could not be identified, we requested one from the authors of the empirical study. If a protocol was not available, we compared the outcomes reported in the methods section of the publication with the outcomes reported in the results section of the publication.

Data synthesis

For each of the five types of selective inclusion and reporting, we combined estimates of the primary outcome of overall prevalence of selective inclusion or reporting (primary outcome #1) in a meta-analysis using a random-effects model. We used DerSimonian and Laird's method of moments estimator to estimate the between trial variance (DerSimonian 1986). When it was not possible to meta-analyse (because similar measures of overall prevalence had not been measured in more than one empirical study or were incompletely reported), we reported percentages and 95% CIs for each empirical study in tables. For each of the five types of selective inclusion and reporting, we combined RR estimates of associations between overall prevalence of selective inclusion or reporting and

statistical significance or direction of effect (primary outcomes #2 and #3) in a meta-analysis using a random-effects model. Where data could not be combined in a meta-analysis, we reported RR estimates, 95% CIs and statistical significance for each empirical study in tables. We did not plan to meta-analyse estimates of bias due to selective inclusion. Instead, we have reported estimates of bias as reported by the empirical study investigators in tables.

Subgroup analysis and investigation of heterogeneity

We planned to undertake subgroup analyses to investigate whether the primary prevalence and association outcomes were modified by whether the systematic reviews included in the empirical studies had a review protocol or not; and whether or not the clinical topic of the systematic reviews had an established core outcome set (COMET Initiative).

Sensitivity analysis

We planned to assess the robustness of meta-analytic effect estimates based on the risk of bias of the included empirical studies. Specifically, we planned to compare meta-analytic effect estimates that included all eligible empirical studies to meta-analytic effect estimates that included empirical studies only if they were rated at low risk of selection bias, and low or unclear risk of selective reporting bias. We used the criterion of low or unclear risk of selective reporting bias because we anticipated that most empirical studies would not have a study protocol and hence would be rated as being at unclear risk of bias on this domain. We also planned to perform sensitivity analyses to determine the robustness of meta-analysed effect estimates based on the different definitions of selective inclusion, discrepant outcome reporting or selective partial reporting used by the empirical study investigators. In cases where ORs were converted to RR, and we were unable to determine the baseline risk in the empirical study, we planned to undertake sensitivity analyses assuming a range of baseline risks from the other included studies.

We undertook a *post-hoc* sensitivity analysis to investigate whether the meta-analytic proportion was robust to the method of analysis. In our primary analysis, we meta-analysed raw proportions. However, other methods have been shown to be preferential because they remove bias that can arise from assuming the normal approximation to the binomial and because of the correlation between the proportion and its variance (Trikalinos 2013). We undertook meta-analyses of logit transformed proportions, arcsine transformed proportions, Freeman-Tukey double arcsine transformed proportions (Freeman 1950) and fitted a random-effects logistic regression (binomial-normal model). These sensitivity analyses were performed using the 'metafor' package in the statistical package 'R' (Viechtbauer 2010).

RESULTS

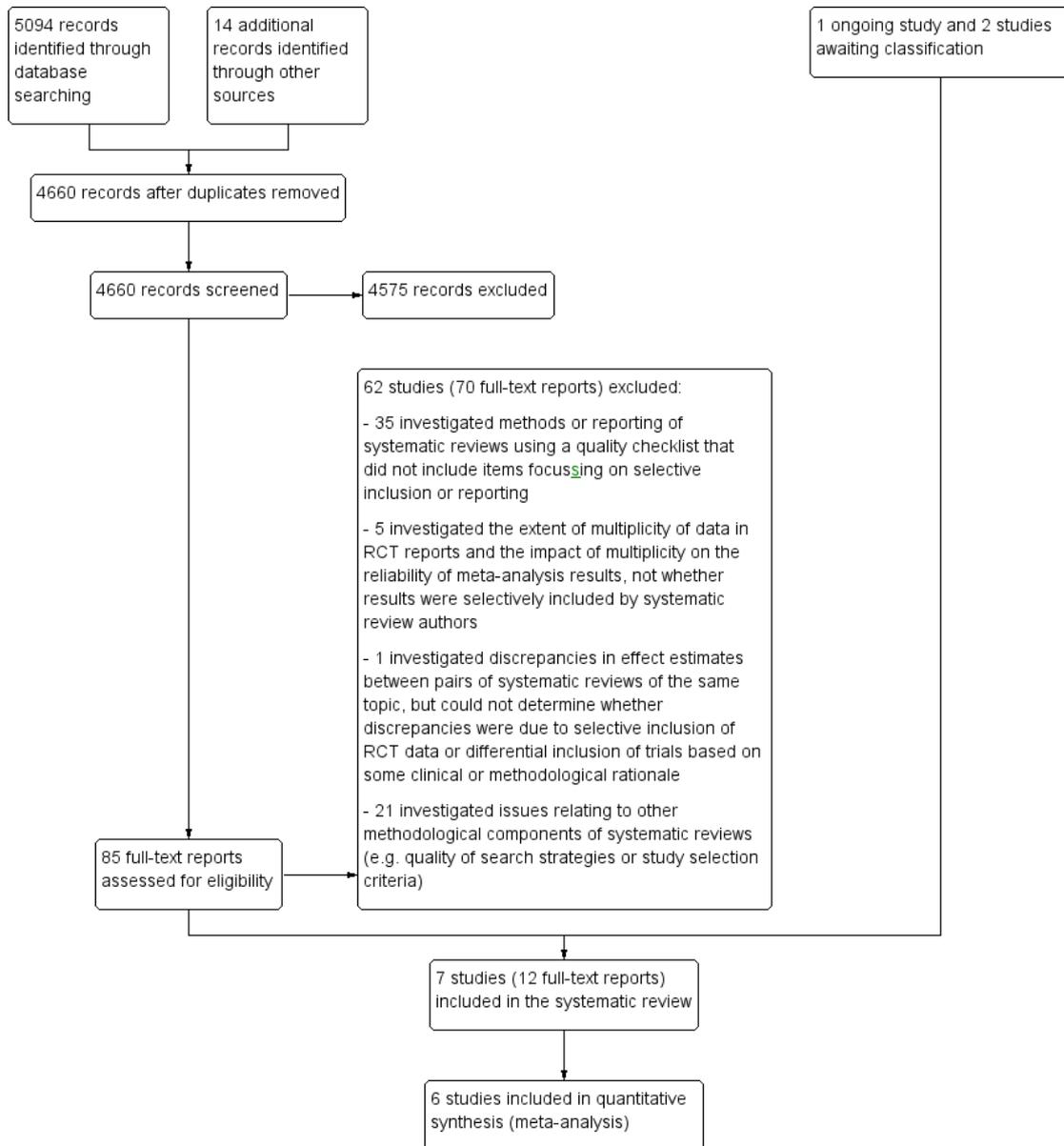
Description of studies

See [Characteristics of included studies](#); [Characteristics of excluded studies](#); [Characteristics of studies awaiting classification](#); and [Characteristics of ongoing studies](#).

Results of the search

The search of electronic databases identified a total of 5094 records. We identified 14 additional records through other sources. After we removed duplicates, 4660 records remained. From screening titles and abstracts, we excluded 4575 records that were not relevant to the review, and retrieved 85 full text reports for further examination. Of these, seven studies fulfilled the inclusion criteria. Two additional studies are awaiting classification as results are only available as a conference abstract and are currently being drafted for full publication ([Johnston 2012](#); [Middleton 2010](#)) and one study is ongoing ([Page 2013b](#)). [Figure 1](#) depicts a flow diagram of the study selection process.

Figure 1. Study flow diagram.



Included studies

See [Characteristics of included studies](#).

The seven included studies were published between 2002 and 2013. Four studies were published in a peer-reviewed journal (Beller 2011; Dwan 2013a; Kirkham 2010b; Silagy 2002) while three were available as conference abstracts only (Hopewell 2010; Parmelli 2007; Vlassov 2008). The median number of included systematic reviews per study was 100 (range: 46 to 288). Six studies only included Cochrane Reviews (Dwan 2013a; Hopewell 2010; Kirkham 2010b; Parmelli 2007; Silagy 2002; Vlassov 2008), while one included both Cochrane and non-Cochrane reviews (Beller 2011). The years of publication of systematic reviews included in the studies ranged from 1996 to 2009. In six studies, various areas of health care were addressed in the included systematic reviews, while one study only included systematic reviews of interventions for cystic fibrosis and genetic disorders (Dwan 2013a). No study assessed the methodological quality of included systematic reviews (e.g. using the Overview Quality Assessment Questionnaire (OQAQ, Oxman 1991) or AMSTAR (Shea 2007)). Only one study recorded the proportion of systematic reviews that only included RCTs (85%) (Dwan 2013a). However, five of the remaining six studies only included Cochrane Reviews (Hopewell 2010; Kirkham 2010b; Parmelli 2007; Silagy 2002; Vlassov 2008), which infrequently include NRS (CEU 2012; Moher 2007). No study assessed the extent of overlap of RCTs in the included systematic reviews. In terms of the types of selective inclusion and reporting investigated, no study investigated selective inclusion of RCT outcome data in systematic reviews, none investigated discrepancies between systematic review registry entries and published systematic reviews, four investigated discrepant outcome

reporting between systematic review protocols and published systematic reviews (Dwan 2013a; Kirkham 2010b; Parmelli 2007; Silagy 2002), two investigated discrepant outcome reporting between the full text and abstract of systematic reviews (Hopewell 2010; Vlassov 2008), and one investigated selective partial reporting in systematic reviews (Beller 2011).

Excluded studies

See [Characteristics of excluded studies](#).

We excluded 62 studies (70 full-text reports). The main reasons for exclusion were: (i) 35 studies investigated the methods or reporting of systematic reviews using a quality checklist that did not include items focusing on selective inclusion or reporting (e.g. OQAQ, AMSTAR, QUOROM or PRISMA); (ii) five studies investigated the extent of multiplicity of data in RCT reports and the impact of multiplicity on the reliability of meta-analysis results, but not whether results were selectively included by systematic review authors; (iii) one study investigated discrepancies in meta-analytic effect estimates between pairs of systematic reviews of the same topic, but could not determine whether discrepancies were due to selective inclusion of RCT data or differential inclusion of trials based on some clinical or methodological rationale; and (iv) 21 studies investigated issues relating to other methodological components of systematic reviews, such as the quality of search strategies or study selection criteria (these 21 studies are not listed in the table of [Characteristics of excluded studies](#) but are available from the review authors on request).

Risk of bias in included studies

See [Figure 2](#).

Figure 2. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.

| | Is the empirical study at low risk of selection bias? | Is the empirical study at low risk of selective reporting bias? |
|---------------|---|--|
| Beller 2011 |  |  |
| Dwan 2013a |  |  |
| Hopewell 2010 |  |  |
| Kirkham 2010b |  |  |
| Parmelli 2007 |  |  |
| Silagy 2002 |  |  |
| Vlassov 2008 |  |  |

We rated all studies as having a low risk of selection bias because either a random sample of systematic reviews or an equivalent sampling process (e.g. systematic sampling) that was unlikely to introduce bias was used, or all systematic reviews registered or published during a particular time period were included in the study. Three studies were rated as having a low risk of selective reporting bias because data for all outcomes and analyses that were pre-specified in an unpublished study protocol were either reported in the publications or provided by the authors (Beller 2011; Dwan 2013a; Kirkham 2010b). The remaining four studies were rated as having an unclear risk of selective reporting bias because it was unclear whether all measured and analysed outcomes were reported (Hopewell 2010; Parmelli 2007; Silagy 2002; Vlassov 2008).

Effect of methods

Selective inclusion of RCT outcome data in systematic reviews

We did not identify any studies.

Discrepancies between systematic review registry entries and published systematic reviews

We did not identify any studies.

Discrepancies between systematic review protocols and published systematic reviews

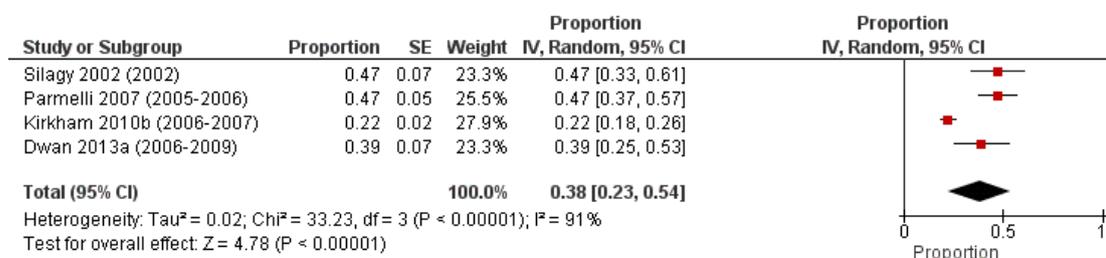
Four studies, including 485 Cochrane Reviews, reported the prevalence of systematic reviews with discrepant outcome reporting between the protocol and published systematic review (Dwan 2013a; Kirkham 2010b; Parmelli 2007; Silagy 2002).

Primary outcomes

Overall prevalence of selective reporting

To estimate the overall prevalence of discrepant outcome reporting, we combined prevalence estimates of systematic reviews with any discrepancy in at least one outcome between the protocol and published systematic review (e.g. adding a new outcome, omitting a protocol-defined outcome, upgrading a secondary outcome to primary or downgrading a primary outcome to secondary). The combined prevalence of systematic reviews that added, omitted, upgraded or downgraded at least one outcome between the protocol and published systematic review was 38% (95% CI 23% to 54%; Figure 3). There was considerable statistical heterogeneity in the prevalence estimates ($I^2 = 91%$, 95% CI 79% to 96%). The reason for this heterogeneity is unclear. Kirkham 2010b had a lower prevalence estimate than the other studies, but we could not identify any variation in the types of reviews included or anything methodologically unique about Kirkham 2010b that could explain the statistical heterogeneity. The studies included Cochrane Reviews published between 2002 and 2009, and one post-hoc explanation for the observed heterogeneity that we considered was whether there had been changes to editorial processes used in Cochrane Reviews over time. However, to the best of our knowledge, the only relevant change to editorial processes was that 'Differences between protocol and review' was introduced as a standard heading in Cochrane Reviews in 2008. This heading may discourage authors from changing outcomes between the protocol and review, so its introduction could have reduced the prevalence of discrepant outcome reporting in Cochrane Reviews published after 2008. However, this change would only effect a small proportion of reviews included in Dwan 2013a and in none of the other studies so it does not explain the observed statistical heterogeneity.

Figure 3. Random-effects meta-analysis of proportion of systematic reviews with any discrepancy in at least one outcome from protocol to published systematic review.



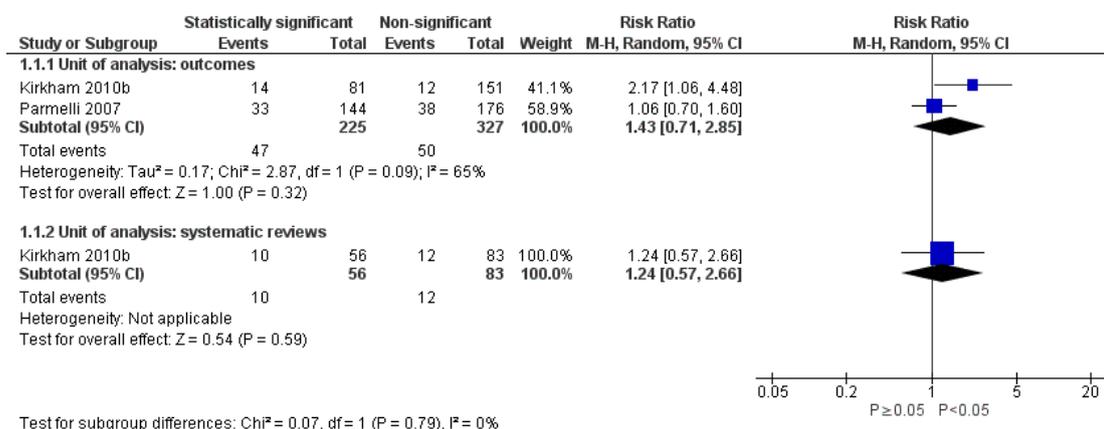
Association between statistical significance and selective reporting

Two studies investigated the association between statistical significance and discrepant outcome reporting (Kirkham 2010b; Parmelli 2007). The authors of Dwan 2013a and Silagy 2002 confirmed that neither study investigated this association. In both Kirkham 2010b and Parmelli 2007, statistical significance of the meta-analysis was defined as $P < 0.05$ and types of discrepancies included adding, upgrading or downgrading outcomes between the protocol and the published systematic review. The unit of analysis in both studies was the outcome, which means that more than one outcome per systematic review may have contributed to

the analysis. Kirkham 2010b also examined the association at the systematic review level (i.e. the number of systematic reviews with at least one discrepant outcome).

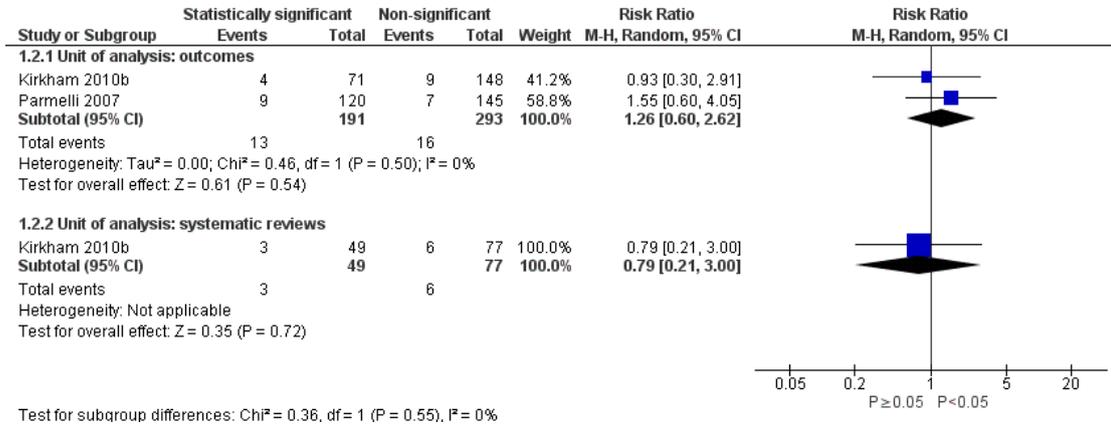
The association between statistical significance and adding or upgrading an outcome between protocol and published systematic review was uncertain. Our meta-analysis of both studies analysed at the outcome level, including 552 meta-analyses, suggested an increased risk of adding or upgrading when the outcome was statistically significant, although the 95% CI included no association and a decreased risk as plausible estimates of association (RR 1.43, 95% CI 0.71 to 2.85; Figure 4; Analysis 1.1). The analysis at the review level in Kirkham 2010b produced a similar effect estimate (RR 1.24, 95% CI 0.57 to 2.66; N = 139 systematic reviews).

Figure 4. Forest plot of association between statistical significance and outcome adding/upgrading.



There was also uncertainty in the association between statistical significance and downgrading an outcome between protocol and published systematic review. Downgrading was less common than adding or upgrading (Figure 5). Our meta-analysis of both studies analysed at the outcome level, including 484 meta-analyses, suggested an increased risk of downgrading when the outcome was statistically significant although the 95% CI included no association and a decreased risk as plausible estimates of association (RR 1.26, 95% CI 0.60 to 2.62; Figure 5; Analysis 1.2). Also, the direction of the estimated association differed for these two studies. The RR when analysed at the review level in Kirkham 2010b was below one, although the 95% CI included no association and an increased risk as plausible estimates of association (RR 0.79, 95% CI 0.21 to 3.00; N = 126 systematic reviews).

Figure 5. Forest plot of association between statistical significance and outcome downgrading.



Association between direction of the effect estimate and selective reporting

No study investigated the association between direction of the effect estimate and discrepant outcome reporting (confirmed via contact with authors of all four studies).

Secondary outcomes

The prevalences of specific types of discrepant outcome reporting (e.g. adding a new outcome, downgrading a primary outcome to secondary) are shown in Table 1.

Table 1. Prevalence of specific types of discrepant outcome reporting between systematic review protocols and published systematic reviews

| Type of discrepant outcome reporting | Silagy 2002 (2002) ¹ N = 47 | Parmelli 2007 (2005 to 2006) ¹ N = 104 | Kirkham 2010b (2006 to 2007) ¹ N = 288 | Dwan 2013a (2006 to 2009) ¹ N = 46 |
|--|---|--|--|--|
| Any discrepancy in at least one outcome (primary, secondary, or unlabelled) from protocol to published systematic review | 22 (47%) | 49 (47%) | 64 (22%) | 18 (39%) |
| Any discrepancy in at least one primary outcome from protocol to published systematic review | -- | -- | 48 (17%) | -- |

(Continued)

| | | | | |
|--|---------|----------|---------|----------|
| Upgrade of at least one outcome from secondary or unlabelled in the protocol to primary in the published systematic review | -- | 11 (11%) | 24 (8%) | 3 (7%) |
| Downgrade of at least one outcome from primary in the protocol to secondary or unlabelled in the published systematic review | -- | 8 (8%) | 12 (4%) | 10 (22%) |
| Addition of at least one new outcome (primary, secondary or unlabelled) in the published systematic review that was not specified in the protocol | 9 (19%) | 32 (31%) | -- | 7 (15%) |
| Addition of at least one new primary outcome in the published systematic review that was not specified in the protocol | -- | -- | 8 (3%) | 3 (7%) |
| Addition of at least one new secondary outcome in the published systematic review that was not specified in the protocol | -- | -- | -- | 4 (9%) |
| Addition of new measurement instruments or criteria for existing outcomes in the published systematic review that were not specified in the protocol | 7 (15%) | -- | -- | -- |
| Omission of at least one outcome (primary, secondary or unlabelled) from the published systematic review which was listed in the protocol | 6 (13%) | 23 (22%) | -- | 5 (11%) |

(Continued)

| | | | | |
|--|----|----|--------|--------|
| Omission of at least one primary outcome from the published systematic review which was listed in the protocol | -- | -- | 7 (2%) | 2 (4%) |
| Omission of at least one secondary outcome from the published systematic review which was listed in the protocol | -- | -- | -- | 3 (7%) |

¹indicates range of years of publication of included systematic reviews.

'--' indicates that the type of discrepancy was not examined by the study authors.

In two studies, the reasons for discrepant outcome reporting were sought in the reports of the published systematic reviews (Dwan 2013a; Kirkham 2010b). Kirkham 2010b reported that only four of 64 (6%) and Dwan 2013a reported that only four of 18 (22%) systematic reviews described the reasons for discrepant outcome reporting. In two studies, the systematic review authors were contacted to seek reasons for the discrepancies (Kirkham 2010b; Silagy 2002). Kirkham 2010b contacted the authors of 48 systematic reviews with any discrepancy in at least one primary outcome, of whom 34 replied but only 28 could recall the reason for the discrepancy. Of these 28 systematic reviews, there was potential for bias in eight (29%) as changes were made to the primary outcome after reading the results of the individual trials. Silagy 2002 contacted authors (65% response rate) to clarify the reasons for discrepancies in any section of the systematic review (e.g. background, search strategy, outcomes). However, we were unable to determine which reasons were relevant to discrepancies in outcomes. Despite this, Silagy 2002 reported that none of respondents' stated reasons for discrepancies appeared to be related to knowledge of the results of individual trials. Parmelli 2007 did not assess reasons for discrepant outcome reporting.

Discrepancies between full text and abstract of systematic reviews

Two studies, including 152 Cochrane Reviews, reported the prevalence of systematic reviews with discrepant outcome reporting be-

tween the full text and abstract (Hopewell 2010; Vlassov 2008).

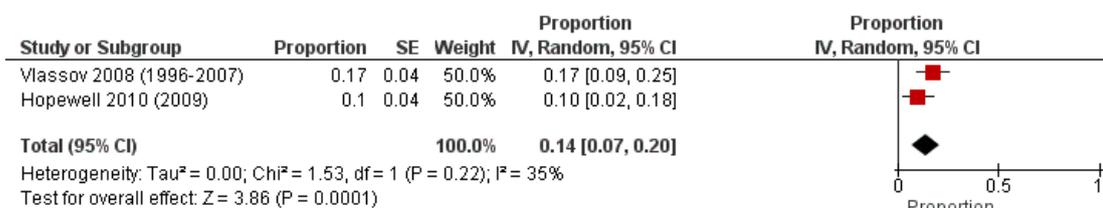
Primary outcomes

None of the primary outcomes for this review were measured in these studies (confirmed via contact with the authors).

Secondary outcomes

One study reported the prevalence of systematic reviews that reported only a subset of primary outcomes from the full text of the review in the abstract (Hopewell 2010). Vlassov 2008 confirmed that he did not assess this outcome. Of 62 Cochrane Reviews, 20 (32%, 95% CI 21% to 45%) did not report all primary outcomes in the abstract, where reporting was defined as presenting an effect estimate or at least stating whether or not the effect estimate was statistically or clinically significant. Of the 20 systematic reviews that did not report all primary outcomes in the abstract, the reason for non-reporting was because either none of the included trials reported data for the outcome (nine reviews) or the reason was unclear (11 reviews). There was insufficient data reported to determine whether there was an association between statistical significance and non-reporting of primary outcomes in the abstract. Both studies reported the prevalence of systematic reviews presenting the results of a secondary outcome before the results of any primary outcomes in the abstract. No overlapping systematic reviews were included in the two studies. The combined prevalence of systematic reviews reporting secondary outcomes before a primary outcome was 14% (95% CI 7% to 20%; Figure 6). Insufficient data was reported to determine whether there was an association between statistical significance and the order in which outcomes were reported.

Figure 6. Random-effects meta-analysis of proportion of systematic reviews presenting the results of a secondary outcome before the results of the primary outcome(s) in the abstract.



Selective partial reporting in systematic reviews

One study, including 182 systematic reviews (64 Cochrane Reviews and 118 non-Cochrane reviews), assessed the completeness of reporting of the primary outcome in the abstracts of systematic reviews (Beller 2011).

Primary outcomes

None of the primary outcomes for this review were measured in this study (confirmed via contact with the author).

Secondary outcomes

The association between statistical significance and complete reporting of primary outcomes in the abstract was reported. The study authors defined complete reporting as meeting the following minimum criteria: “The direction and size of effect can be determined from the wording or deduced numerically, and a measure of precision of the effect size was stated (e.g. CI)”. Statistically significant primary outcomes were more likely to be completely reported in the abstract than non-significant primary outcomes (85/112 versus 20/70; RR 2.66, 95% CI 1.81 to 3.90).

Subgroup and sensitivity analyses

There were a few overlapping systematic reviews across the Dwan 2013a and Kirkham 2010b studies. In a sensitivity analysis, removal of Dwan 2013a (which included fewer systematic reviews) resulted in the same meta-analytic estimate of the prevalence of systematic reviews with at least one discrepant outcome between the protocol and published systematic review (38%, 95% CI 18% to 58%; N = 439 systematic reviews). We were unable to undertake any of our other planned subgroup analyses because there was either an insufficient number of studies or insufficient data available.

Our post-hoc sensitivity analysis that investigated the impact of combining proportions using a different transformation (logistic),

variance stabilising functions (arcsine and double arcsine) and random effects logistic regression found that the meta-analytic estimate was robust to the method used. The meta-analytic proportion and its CI was 0.38 (95% CI 0.24 to 0.54) using the logit transformation; 0.38 (95% CI 0.23 to 0.54) using the arcsine transformation; 0.38 (95% CI 0.23 to 0.54) using the double arcsine transformation; and 0.37 (95% CI 0.26 to 0.49) in the random-effects logistic regression.

DISCUSSION

Summary of main results

Based on a meta-analysis of four studies, 38% (95% CI 23% to 54%) of systematic reviews added, omitted, upgraded or downgraded at least one outcome between the protocol and published systematic review. The association between statistical significance and discrepant outcome reporting was uncertain. Two studies found that reasons for discrepant outcome reporting were infrequently reported in published systematic reviews. One study found that 32% (95% CI 21% to 45%) of systematic reviews did not report all primary outcomes in the abstract. Another study found that statistically significant primary outcomes were more likely to be completely reported in the systematic review abstract than non-significant primary outcomes (RR 2.66, 95% CI 1.81 to 3.90).

Overall completeness and applicability of evidence

Some of the types of selective inclusion or reporting which we specified in our protocol have not been empirically investigated. No study has investigated selective inclusion of RCT outcome data in systematic reviews, although one such study is currently underway by authors of this review (Page 2013b). This study will explore whether, in the presence of multiplicity of outcome data, trial effect estimates selected for inclusion in a sample of meta-

analyses are systematically more or less favourable to the intervention than what would be expected under a process consistent with random selection. No study has investigated discrepancies between systematic review registry entries and published systematic reviews, perhaps because a publicly available register for systematic review protocols was only launched in February 2011 (Booth 2013). Also, selective partial reporting of outcomes has only been evaluated in abstracts, not in the full text of systematic reviews.

No study had investigated another type of selective reporting that we did not define in our protocol, but identified through other work (Page 2013a). This is the practice where systematic review authors select a meta-analytic effect estimate to report after conducting multiple meta-analyses of the different data available in trial reports for a particular outcome. For example, meta-analyses of pain could include final values only, change from baseline values only or a mixture of these two, and systematic review authors may choose to report one of these meta-analytic effects based on the results. One way to investigate this type of selective reporting of analyses would be to compare the analysis plans reported in the systematic review protocol to the analyses reported in the published systematic review. Similar studies have been conducted at the RCT level (Al-Marzouki 2008; Chan 2008; Vedula 2009; Vedula 2013). These types of investigations are strengthened by retrieval of the studies included in each systematic review to determine whether the systematic review authors could have analysed the RCT data in multiple ways.

None of the included studies evaluated systematic reviews published after 2009. Several initiatives occurred around 2009 that may have reduced the prevalence of selective reporting in more recent reviews. For example, a 'Differences between protocol and review' heading was introduced as a standard heading in Cochrane Reviews in 2008. This heading may discourage authors from changing outcomes between the protocol and review unless they have a good reason to do so. In July 2009, the PRISMA Statement appeared (updating the QUOROM statement), which advises that changes to the outcomes made after the review has started should be described and that justification should be provided. In addition, the PRISMA Statement advises that when available, summary data per group, effect estimates and CIs should be reported for all outcomes considered in the systematic review, regardless of their results (Liberati 2009; Moher 2009). The PRISMA for Abstracts Statement (published April 2013) advises authors to clearly indicate in the abstract the protocol-defined, pre-specified importance of each outcome (i.e. primary or secondary) and to not only report outcomes with statistically or clinically significant results (Beller 2013). Similar recommendations are provided in the MECIR reporting standards (published in December 2012), which summarise attributes of reporting considered either mandatory (compliance required for publication) or highly desirable (expected but may be justifiably not done) for Cochrane intervention reviews (Chandler 2012). Given that improvements have been seen in the quality of reporting of RCTs following journal endorsement of the

CONSORT reporting guideline (Turner 2012), it might be expected that similar improvements will occur for systematic reviews. The four studies investigating discrepant outcome reporting between systematic review protocols and published systematic reviews only focused on Cochrane Reviews. This limits the generalisability of the findings, as Cochrane Reviews comprise a minority of all published systematic reviews (Moher 2007). It is not surprising that discrepant outcome reporting has not been investigated in non-Cochrane reviews since only 10% of non-Cochrane reviews explicitly report working from a review protocol, with 4% working from a publicly available protocol (Turner 2013). The potential for selective reporting is possibly higher in such reviews, for similar reasons outlined in regards to selective reporting in NRS (Norris 2013). First, systematic review authors not working from a protocol may approach the selection of outcomes to include in an iterative manner, changing their position about the importance of particular outcomes while they identify eligible studies, analyse the data and draft the manuscript. In this scenario, the results of included studies, coupled with existing beliefs about the effectiveness of an intervention, may bias the decisions about the final set of outcomes to report in the review (rather than adhering to a pre-specified clinical rationale guiding such decisions). Second, review authors who pre-specify outcomes in a protocol may feel less obliged to follow it if it is not publicly available (Stewart 2012). Protocols for systematic reviews conducted outside of organisations such as The Cochrane Collaboration are becoming increasingly available with the launch of the PROSPERO database of prospectively registered systematic reviews (Booth 2013), a new BioMed Central journal (*Systematic Reviews*) for systematic reviews and associated research (Moher 2012), and the upcoming PRISMA for Protocols (PRISMA-P) reporting guideline (Shamseer 2013). These initiatives will allow for future investigations of selective inclusion and reporting in non-Cochrane reviews.

An unanticipated finding in our review was that the meta-analytic risk of downgrading an outcome between protocol and published systematic review was increased when the meta-analysis result for the downgraded outcome was statistically significant. This is in contrast to the association previously found at the RCT level, where outcomes were more likely to be downgraded from trial registry entry to published RCT if the result was non-significant (Mathieu 2009). While this observed association may be a result of chance, it may be the case that the impetus for upgrading and downgrading outcomes differs depending on the direction of effect, type of outcome and the review authors' preconceptions about the effectiveness and safety of the interventions. For example, in systematic reviews comparing active interventions with placebo, motivation to find an effective, safe intervention may be greater, thus statistically significant efficacy outcomes which favour the active intervention may be more likely to be upgraded, while statistically significant safety outcomes that favour the placebo may be more likely to be downgraded. However, in systematic reviews that compare two active interventions, there may be motivation

to both upgrade and downgrade outcomes if systematic review authors have preconceptions about the effectiveness and safety of the included interventions. Neither [Kirkham 2010b](#) or [Parmelli 2007](#) reported the percentage of systematic reviews by type of outcome or type of comparison, nor did they examine whether the association was modified by these two factors (confirmed via personal communication). Further investigation of these issues would be useful.

None of the included studies reported sufficient data to enable investigation of whether the prevalence and association between statistical significance and discrepant outcome reporting were modified by the use of a core outcome set, which is a standardised set of outcomes to measure in all clinical trials and include in all systematic reviews of a particular condition ([Williamson 2012a](#); [Williamson 2012b](#)). The use of a core outcome set would be expected to reduce bias in systematic reviews. A 2012 survey of Coordinating Editors of Cochrane Review Groups showed that 36% of the groups have a centralised policy regarding which outcomes to include in the 'Summary of Findings' table, which is a standardised table presenting results of up to seven of the most important outcomes of the review ([Kirkham 2013](#)).

Quality of the evidence

In studies investigating discrepant outcome reporting between the protocol and published systematic review ([Dwan 2013a](#); [Kirkham 2010b](#); [Parmelli 2007](#); [Silagy 2002](#)), methodological weaknesses in the empirical studies exist. The risk of bias was rated as low for both domains in two studies ([Dwan 2013a](#); [Kirkham 2010b](#)), while two studies had a low risk of selection bias but an unclear risk of selective reporting bias ([Parmelli 2007](#); [Silagy 2002](#)). The 95% CI of the combined prevalence estimate was wide, although all estimates within the confidence limits were of concern. The studies could have benefited from examining if the association between statistical significance and changing the status of an outcome (adding, upgrading or downgrading) was modified by the type of comparison, direction of effect and type of outcome (efficacy or safety). Some studies missed opportunities for analysis, e.g. two studies did not investigate the factors associated with the discrepant outcome reporting ([Dwan 2013a](#); [Silagy 2002](#)), and two studies did not assess reasons for discrepant outcome reporting ([Parmelli 2007](#); [Silagy 2002](#)). These additions would have helped to disentangle bias-related reasons from non-bias-related reasons (e.g. omitting a protocol-defined outcome from the published systematic review because no RCTs measured the outcome ([Liberati 2009](#); [Stewart 2012](#))).

Methodological weaknesses also exist in studies investigating selective reporting practices in systematic review abstracts ([Beller 2011](#); [Hopewell 2010](#); [Vlassov 2008](#)). The risk of bias was rated as low for both domains in one study ([Beller 2011](#)), while the other two studies had a low risk of selection bias but an unclear risk of selective reporting bias ([Hopewell 2010](#); [Vlassov 2008](#)). Neither of the

two studies investigating discrepant outcome reporting between the full text and abstract formally investigated factors associated with discrepant outcome reporting (e.g. statistical significance or direction of effect).

Potential biases in the review process

There are some limitations to our methods. While we believe that all relevant, published studies were identified, we have no way of verifying this because a validated search strategy for locating empirical studies of selective inclusion and reporting does not exist. As outlined under [Differences between protocol and review](#), we modified the search strategies that we had specified in our protocol for Ovid MEDLINE, Ovid EMBASE, Ovid PsycInfo and ISI Web of Science because the original search strategies yielded an unexpectedly large number of citations for systematic reviews of intervention studies. These citations were most likely retrieved because terms such as 'bias' or 'selective reporting' were mentioned in the abstract of the systematic review. We therefore modified our original search strategies to a similar one used in [Dwan 2011](#), and these updated strategies were reviewed by an information specialist. We also supplemented our original search strategies with searches of three sources that are comprised solely of records of methodological work in research synthesis. In addition, we were able to obtain and use unpublished data from four studies ([Beller 2011](#); [Hopewell 2010](#); [Parmelli 2007](#); [Vlassov 2008](#)), which is likely to have reduced the chance of publication bias impacting on the results of our review.

Two of us (JK and KD) were authors of two of the included studies ([Dwan 2013a](#); [Kirkham 2010b](#)). Neither author was involved in the data extraction or risk of bias assessment for these two studies. We could have conducted risk of bias assessments under blinded conditions (i.e. where the assessor is unaware of the study author's name, institution, sponsorship, journal, etc.) but evidence regarding whether this process yields systematically different assessments is inconsistent ([Morissette 2011](#)).

While we planned to synthesise results of empirical studies including only systematic reviews of RCTs, we found that most studies did not record whether the included systematic reviews included RCTs only. However, since most of the evidence in this review is based on Cochrane Reviews, and Cochrane Reviews infrequently included NRS prior to 2012 ([CEU 2012](#); [Moher 2007](#)), we believe our results are unlikely to be affected in an important way by the inclusion of some systematic reviews of NRS.

As outlined under [Differences between protocol and review](#), we re-labelled one of the five types of selective inclusion and reporting which we planned to focus on (from 'partial reporting in systematic reviews' to 'selective partial reporting in systematic reviews') while working on the review. This is because we identified 13 studies that evaluated systematic reviews using a reporting quality checklist, such as the PRISMA Statement, during the study screening for this review ([Assendelft 1995](#); [Auperin 1997](#); [Aytug 2012](#); [CEU](#)

2011; CEU 2012; Faggion Jr 2014; Gianola 2013; Ma 2012; Minozzi 2006a; Minozzi 2006b; Moseley 2009; Roundtree 2009; Shea 2006). These 13 studies collected data on whether outcomes were partially reported (e.g. reporting an effect estimate without a measure of variation). However, the authors defined the reporting practices they evaluated as markers of “reporting quality”, not “selective reporting”, and none attempted to investigate whether the reasons for partial reporting were related to the results (i.e. selective partial reporting). Studies evaluating “reporting quality” of systematic reviews fall outside the scope of our review.

Agreements and disagreements with other studies or reviews

A 2008 systematic review that synthesised results of methodology reviews investigating any type of bias that can occur throughout the systematic review process did not identify any reviews examining biases associated with selecting studies for inclusion, selecting outcomes, synthesising studies or reporting outcomes (Tricco 2008). To the best of our knowledge, no subsequent systematic reviews of studies investigating such practices have been published. The most similar systematic review to ours is a methodology review of studies examining discrepant outcome reporting at the RCT level (Dwan 2011). This review identified six studies comparing outcomes between protocols and published reports of RCTs, and three studies comparing outcomes between trial registry entries and published reports of RCTs. The studies examined heterogeneous samples of RCTs and were not combined in a meta-analysis. The authors reported that at least one primary outcome was added, omitted, upgraded or downgraded in 4% to 50% of RCT publications. The equivalent outcome in our review had a narrower range of prevalence estimates (i.e. 22% to 47%). In another systematic review, statistical significance was associated with complete reporting of outcomes in the full text of RCTs (Dwan 2013b). The magnitude of this association was similar to the association between statistical significance and complete reporting of the primary outcome in the systematic review abstract, as reported in Beller 2011.

AUTHORS' CONCLUSIONS

Implication for methodological research

Potential future studies could investigate: (1) potential selective inclusion of results in systematic reviews; (2) discrepant outcome reporting between registry entries and published systematic reviews (e.g. using registry entries in the PROSPERO database of prospectively registered systematic reviews); (3) discrepant outcome reporting between non-Cochrane systematic review protocols and published systematic reviews (e.g. using protocols published in journals such as Systematic Reviews); (4) whether discrepant outcome reporting between protocols and published systematic reviews has reduced since the publication of the PRISMA Statement; (5) whether the prevalence of selective partial reporting in systematic review abstracts has reduced following publication of the PRISMA for Abstracts Statement; and (6) whether selective inclusion and reporting occurs less frequently in systematic reviews with a core outcome set. Where possible, investigators should assess both the prevalence and factors associated with selective inclusion and reporting in systematic reviews, and whether the associations are modified by the type of comparison, direction of effect and type of outcome.

ACKNOWLEDGEMENTS

We thank Mike Clarke for his support, advice and assistance throughout the conception and preparation of this review. We acknowledge Sarah Chapman for her help with developing the search strategies. Also we are grateful to Elaine Beller, Sally Hopewell, Toby Lasserson, Elena Parmelli and Vasya Vlassov for providing us with unpublished study data. We thank Asbjørn Hróbjartsson and Isabelle Boutron for their helpful comments, and Peter Gøtzsche for his helpful editing of the review.

REFERENCES

References to studies included in this review

Beller 2011 *{published and unpublished data}*

Beller E, Glasziou P, Hopewell S, Altman D. Reporting of effect direction and size in abstracts of systematic reviews (abstract). Oral presentation at the 19th Cochrane Colloquium; 2011 Oct 19-22; Madrid, Spain. 2011. 2011: 8.

* Beller EM, Glasziou PP, Hopewell S, Altman DG. Reporting of effect direction and size in abstracts of systematic reviews. *JAMA* 2011;**306**(18):1981–2.

Dwan 2013a *{published and unpublished data}*

* Dwan K, Kirkham JJ, Williamson PR, Gamble C. Selective reporting of outcomes in randomised controlled trials in systematic reviews of cystic fibrosis. *BMJ Open* 2013;**3**(6):e002709.

Dwan K, Williamson P, Gamble C, Remington T, Jahnke N, Kirkham J. Investigating outcome reporting bias in Cochrane Cystic Fibrosis and Genetic Disorder (CFGD) reviews (abstract). Oral presentation at the Joint Cochrane and Campbell Colloquium; 2010 Oct 18-22; Keystone,

Colorado, USA 2010:24.

Hopewell 2010 {published and unpublished data}

Hopewell S, Beller E. Is there any evidence of selective reporting of outcomes in abstracts of Cochrane reviews? (abstract). Oral presentation at the Joint Cochrane and Campbell Colloquium; 2010 Oct 18-22; Keystone, Colorado, USA. 2010:24-5.

Kirkham 2010b {published and unpublished data}

Kirkham J, Altman D, Williamson P. ORBIT study: outcome reporting bias in trials - primary outcomes in Cochrane systematic reviews (abstract). Oral presentation at the 17th Cochrane Colloquium; 2009 Oct 11-14, Singapore. 2009:6.

* Kirkham JJ, Altman DG, Williamson PR. Bias due to changes in specified outcomes during the systematic review process. *PLoS ONE* 2010;5(3):e9810.

Parmelli 2007 {published and unpublished data}

Parmelli E, D'Amico R, Minozzi S, Bassi C, Liberati A. Were the outcomes reported in systematic reviews stated in protocols? A systematic review (abstract). XIV Cochrane Colloquium; 2006 October 23-26; Dublin, Ireland. 2006: 142.

* Parmelli E, Liberati A, D'Amico R. Reporting of outcomes in systematic reviews: comparison of protocols and published systematic reviews (abstract). XV Cochrane Colloquium; 2007 Oct 23-27; São Paulo, Brazil. 2007: 118-9.

Silagy 2002 {published data only}

Silagy C, Middleton P, Hopewell S. Is publishing Cochrane protocols a way to reduce or introduce bias? (abstract). 9th Annual Cochrane Colloquium; 2001 Oct 9-13; Lyon, France. 2001:42.

* Silagy CA, Middleton P, Hopewell S. Publishing protocols of systematic reviews: comparing what was done to what was planned. *JAMA* 2002;287(21):2831-4.

Vlassov 2008 {published and unpublished data}

Vlassov V. Low quality of reporting of primary outcomes in Cochrane abstracts. Poster presentation at the 16th Cochrane Colloquium: Evidence in the era of globalisation; 2008 Oct 3-7; Freiburg, Germany (abstract). *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* 2008;102(Suppl VI):47.

References to studies excluded from this review

Assendelft 1995 {published data only}

Assendelft WJJ, Koes BW, Knipschild PG, Bouter LM. The relationship between methodological quality and conclusions in reviews of spinal manipulation. *JAMA* 1995; 274(24):1942-8.

Auperin 1997 {published data only}

* Auperin A, Pignon J-P, Poynard T. Review article: critical review of meta-analyses of randomized clinical trials in hepatogastroenterology. *Alimentary Pharmacology and Therapeutics* 1997;11(2):215-25.

Auperin A, Pignon JP, Poynard T. Critical review of meta-analyses of randomized clinical trials in

hepatogastroenterology (abstract). Second International Conference Scientific Basis of Health Services & Fifth Annual Cochrane Colloquium; 1997 Oct 8-12; Amsterdam, The Netherlands. 1997:272.

Aytug 2012 {published data only}

Aytug ZG, Rothstein HR, Zhou W, Kern MC. Revealed or concealed? Transparency of procedures, decisions, and judgment calls in meta-analyses. *Organizational Research Methods* 2012;15(1):103-33.

Bhandari 2001 {published data only}

Bhandari M, Morrow F, Kulkarni AV, Tornetta P 3rd. Meta-analyses in orthopaedic surgery. A systematic review of their methodologies. *Journal of Bone and Joint Surgery* 2001;83A(1):15-24.

Bjordal 2003 {published data only}

Bjordal JM. A quantitative study of bias in systematic reviews. *Advances in Physiotherapy* 2003;5(2):83-96.

Bjordal 2005 {published data only}

Bjordal JM, Bogen B, Lopes-Martins RAB, Klovning A. Can Cochrane Reviews in controversial areas be biased? A sensitivity analysis based on the protocol of a Systematic Cochrane Review on low-level laser therapy in osteoarthritis. *Photomedicine and Laser Surgery* 2005;23(5):453-8.

Bohlius 2005 {published data only}

Bohlius J, Weingart O, Trelle S, Engert A. Disentangling the data: variations in data submissions from different players and their potential impact on a systematic review (abstract). XIII Cochrane Colloquium; 2005 Oct 22-26; Melbourne, Australia. 2005:60.

Bow 2010 {published data only}

Bow S, Klassen J, Chisholm A, Tjosvold L, Thomson D, Klassen TP, et al. A descriptive analysis of child-relevant systematic reviews in the Cochrane Database of Systematic Reviews. *BMC Pediatrics* 2010;10:34.

CEU 2011 {published and unpublished data}

Cochrane Editorial Unit. Audit of the abstract, plain language summary and summary of findings tables in published Cochrane Reviews. <http://www.editorial-unit.cochrane.org/audit-abstracts-plain-language-summaries-and-summary-findings-tables> (accessed 27th May 2013).

CEU 2012 {published and unpublished data}

Cochrane Editorial Unit. The Cochrane Library - Revolution or Evolution. Background paper for The Cochrane Collaboration's Strategic Session, Paris, France, 18 April 2012. [www.editorial-unit.cochrane.org/sites/editorial-unit.cochrane.org/files/uploads/2012-CC-strategic-session`meeting-report.pdf](http://www.editorial-unit.cochrane.org/sites/editorial-unit.cochrane.org/files/uploads/2012-CC-strategic-session%20meeting-report.pdf) (accessed 27th May 2013).

Choi 2001 {published data only}

Choi PT, Halpern SH, Malik N, Jadad AR, Tramèr MR, Walder B. Examining the evidence in anesthesia literature: a critical appraisal of systematic reviews. *Anesthesia and Analgesia* 2001;92(3):700-9.

da Costa 2013 {published data only}

da Costa BR, Nüesch E, Rutjes AW, Johnston BC, Reichenbach S, Trelle S, et al. Combining follow-up

and change data is valid in meta-analyses of continuous outcomes: a meta-epidemiological study. *Journal of Clinical Epidemiology* 2013;**66**(8):847–55.

Delaney 2005 {published data only}

Delaney A, Bagshaw SM, Ferland A, Manns B, Laupland KB, Doig CJ. A systematic evaluation of the quality of meta-analyses in the critical care literature. *Critical Care* 2005;**9**(5):R575–82.

Dundar 2006 {published data only}

* Dundar Y, Dodd S, Dickson R, Walley T, Haycox A, Williamson PR. Comparison of conference abstracts and presentations with full-text articles in the health technology assessments of rapidly evolving technologies. *Health Technology Assessment* 2006;**10**(5):iii-iv, ix--145.
Dundar Y, Dodd S, Williamson P, Dickson R, Walley T. Case study of the comparison of data from conference abstracts and full-text articles in health technology assessment of rapidly evolving technologies: does it make a difference?. *International Journal of Technology Assessment in Health Care* 2006;**22**(3):288–94.

Faggion Jr 2014 {published data only}

Faggion CM Jr, Liu J, Huda F, Atieh M. Assessment of the quality of reporting in abstracts of systematic reviews with meta-analyses in periodontology and implant dentistry. *Journal of Periodontal Research* 2014;**49**(2):137–42.

Farmer 2012 {published data only}

Farmer SE, Wood D, Swain ID, Pandyan AD. Assessment of the risk of bias in rehabilitation reviews. *International Journal of Rehabilitation Research* 2012;**35**(4):317–22.

Fishbain 2000 {published data only}

Fishbain D, Cutler RB, Rosomoff HL, Rosomoff RS. What is the quality of the implemented meta-analytic procedures in chronic pain treatment meta-analyses?. *Clinical Journal of Pain* 2000;**16**(1):73–85.

Gianola 2013 {published data only}

Gianola S, Gasparini M, Agostini M, Castellini G, Corbetta D, Gozzer P, et al. Survey of the reporting characteristics of systematic reviews in rehabilitation. *Physical Therapy* 2013;**93**(11):1456–66.

Hartling 2004 {published data only}

Hartling L, Klassen T, Moher D, Tubman M, Chiu A, Wiebe N. Quality of reporting of systematic reviews and its affect on estimates of intervention effectiveness (abstract). 12th Cochrane Colloquium: Bridging the Gaps; 2004 Oct 2-6; Ottawa, Ontario, Canada. 2004:148–9.

Hopewell 2008 {published data only}

Hopewell S, Wolfenden L, Clarke M. Reporting of adverse events in systematic reviews (abstract). XIV Cochrane Colloquium; 2006 October 23-26; Dublin, Ireland. 2006: 48.

* Hopewell S, Wolfenden L, Clarke M. Reporting of adverse events in systematic reviews can be improved: survey results. *Journal of Clinical Epidemiology* 2008;**61**(6):597–602.

Jørgensen 2006 {published data only}

Jørgensen A, Gøtzsche P. Sponsorship, bias and methodology: Cochrane reviews compared with paper-

based meta-analyses of the same drugs (abstract). 12th Cochrane Colloquium: Bridging the Gaps; 2004 Oct 2-6; Ottawa, Ontario, Canada. 2004:80–1.

* Jørgensen AW, Hilden J, Gøtzsche PC. Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review. *BMJ* 2006;**333**(7572):782.

Kelly 2001 {published data only}

Kelly KD, Travers A, Dorgan M, Slater L, Rowe BH. Evaluating the quality of systematic reviews in the emergency medicine literature. *Annals of Emergency Medicine* 2001;**38**(5):518–26.

Kuukasjärvi 2006a {published data only}

Kuukasjärvi P, Malmivaara A, Halinen M, Hartikainen J, Keto PE, Talvensaari T, et al. Overview of systematic reviews on invasive treatment of stable coronary artery disease. *International Journal of Technology Assessment in Health Care* 2006;**22**(2):219–34.

Kuukasjärvi 2006b {published data only}

Kuukasjärvi P, Nordhausen K, Malmivaara A. Reanalysis of systematic reviews: the case of invasive strategies for acute coronary syndromes. *International Journal of Technology Assessment in Health Care* 2006;**22**(4):484–96.

Lacasse 1999 {published data only}

Lacasse Y, Goldstein RS. Overviews of respiratory rehabilitation in chronic obstructive pulmonary disease. *Monaldi Archives for Chest Disease* 1999;**54**(2):163–7.

Ma 2012 {published data only}

Ma B, Qi GQ, Lin XT, Wang T, Chen ZM, Yang KH. Epidemiology, quality, and reporting characteristics of systematic reviews of acupuncture interventions published in Chinese journals. *Journal of Alternative and Complementary Medicine* 2012;**18**(9):813–7.

Minozzi 2006a {published data only}

Minozzi S, Davoli M, Amato L, Vecchi S. Quality of systematic reviews of the Cochrane Drugs and Alcohol Group: can we improve it? (abstract). XIV Cochrane Colloquium; 2006 October 23-26; Dublin, Ireland. 2006.

Minozzi 2006b {published data only}

Minozzi S, Filippini G, Coco L. Quality of the systematic reviews of the Cochrane Multiple Sclerosis Group (abstract). XIV Cochrane Colloquium; 2006 October 23-26; Dublin, Ireland. 2006.

Moher 2007 {published data only}

Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. *PLoS Medicine* 2007;**4**(3):e78.

Moseley 2009 {published data only}

Moseley AM, Elkins MR, Herbert RD, Maher CG, Sherrington C. Cochrane reviews used more rigorous methods than non-Cochrane reviews: survey of systematic reviews in physiotherapy. *Journal of Clinical Epidemiology* 2009;**62**(10):1021–30.

Roundtree 2009 {published data only}

Roundtree AK, Kallen MA, Lopez-Olivo MA, Kimmel B, Skidmore B, Ortiz Z, et al. Poor reporting of search

strategy and conflict of interest in over 250 narrative and systematic reviews of two biologic agents in arthritis: a systematic review. *Journal of Clinical Epidemiology* 2009;**62**(2):128–37.

Sacks 1987 {published data only}

Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomised controlled trials. *New England Journal of Medicine* 1987;**316**(8):450–5.

Sacks 1992 {published data only}

Sacks HS, Berrier J, Reitman D, Pagano D, Chalmers TC. Meta-analysis of randomised controlled trials: an update. In: Balder WC 3rd, Mostellar F editor(s). *Medical Uses of Statistics*. 2nd Edition. Boston, MA: NEJM Books, 1992: 427–442.

Sacks 1996 {published data only}

Sacks HS, Reitman D, Pagano D, Kupelnick B. Meta-analysis: an update. *Mount Sinai Journal of Medicine* 1996; **63**(3-4):216–24.

Schwarzer 2001 {published data only}

Schwarzer G, Antes G, Tallon D, Egger M. Review publication bias? Matched comparative study of Cochrane and journal meta-analyses. 9th Annual Cochrane Colloquium; 2001 Oct 9-13; Lyon, France. 2001.

Shea 2006 {published data only}

Shea B, Bouter LM, Grimshaw JM, Francis D, Ortiz Z, Wells GA, et al. Scope for improvement in the quality of reporting of systematic reviews. From the Cochrane Musculoskeletal Group. *Journal of Rheumatology* 2006;**33**(1):9–15.

Sheikh 2007 {published data only}

Sheikh L, Johnston S, Thangaratinam S, Kilby MD, Khan KS. A review of the methodological features of systematic reviews in maternal medicine. *BMC Medicine* 2007;**5**:10.

Song 1997 {published data only}

Song F, Landes DP, Glenny AM, Sheldon TA. Prophylactic removal of impacted third molars: an assessment of published reviews. *British Dental Journal* 1997;**182**(9): 339–46.

Stroup 2001 {published data only}

* Stroup DF, Thacker SB, Olson CM, Glass RM, Hutwagner L. Characteristics of meta-analyses related to acceptance for publication in a medical journal. *Journal of Clinical Epidemiology* 2001;**54**(7):655–60.

Thacker SB, Stroup DF, Olsen C. Characteristics of meta-analyses submitted to a general medical journal (abstract). Sixth International Cochrane Colloquium; 1998 Oct 22-26; Baltimore, MD, USA. 1998.

Tendal 2011 {published data only}

Tendal B, Nüesch E, Higgins JPT, Jüni P, Göttsche PC. Multiplicity of data in trial reports and the reliability of meta-analyses: empirical study. *BMJ* 2011;**343**:d4829.

Wee 2008 {published data only}

Wee B, Hadley G, Derry S. How useful are systematic reviews for informing palliative care practice? Survey of 25 Cochrane systematic reviews. *BMC Palliative Care* 2008;**7**: 13.

References to studies awaiting assessment

Johnston 2012 {published data only}

Johnston B, Alonso-Coello P, Neumann I, Carrasco-Labra A, Brignardello-Petersen R, Sun X, et al. Reporting of absolute estimates of effect of patient important benefits and harms in abstracts of systematic reviews (abstract). Poster presentation at the 20th Cochrane Colloquium; 2012 Sept 30-Oct 3; Auckland, New Zealand. 2012.

Middleton 2010 {published and unpublished data}

Middleton P. Using Cochrane reviews to help reduce fetal and other perinatal deaths in high income countries (abstract). Oral presentation at the Joint Cochrane and Campbell Colloquium; 2010 Oct 18-22; Keystone, Colorado, USA. 2010.

References to ongoing studies

Page 2013b {published data only}

Page MJ, McKenzie JE, Green SE, Forbes AB. An empirical investigation of the potential impact of selective inclusion of results in systematic reviews of interventions: study protocol. *Systematic Reviews* 2013;**2**:21.

Additional references

Al-Marzouki 2008

Al-Marzouki S, Roberts I, Evans S, Marshall T. Selective reporting in clinical trials: analysis of trial protocols accepted by The Lancet. *Lancet* 2008;**372**(9634):201.

Begg 1988

Begg CB, Berlin JA. Publication bias: a problem in interpreting medical data. *Journal of the Royal Statistical Society, Series A* 1988;**151**(3):419–63.

Beller 2013

Beller EM, Glasziou PP, Altman DG, Hopewell S, Bastian H, Chalmers I, et al. PRISMA for abstracts: reporting systematic reviews in journal and conference abstracts. *PLoS Medicine* 2013;**10**(4):e1001419.

Bender 2008

Bender R, Bunce C, Clarke M, Gates S, Lange S, Pace NL, et al. Attention should be given to multiplicity issues in systematic reviews. *Journal of Clinical Epidemiology* 2008; **61**(9):857–65.

Booth 2010

Booth A, Clarke M, Ghersi D, Moher D, Petticrew M, Stewart L. An international registry of systematic-review protocols. *Lancet* 2010;**377**(9760):108–9.

Booth 2011

Booth A, Clarke M, Ghersi D, Moher D, Petticrew M, Stewart L. Establishing a minimum dataset for prospective registration of systematic reviews: an international consultation. *PLoS One* 2011;**6**(11):e27319.

Booth 2013

Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, et al. PROSPERO at one year: an evaluation of its utility. *Systematic Reviews* 2013;**2**:4.

Chalmers 1990

Chalmers I. Underreporting research is scientific misconduct. *JAMA* 1990;**263**(10):1405–8.

Chan 2004a

Chan AW, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;**291**(20):2457–65.

Chan 2004b

Chan AW, Krleza-Jerić K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ* 2004;**171**(7):735–40.

Chan 2005

Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ* 2005;**330**(7494):753.

Chan 2008

Chan AW, Hróbjartsson A, Jørgensen KJ, Gøtzsche PC, Altman DG. Discrepancies in sample size calculations and data analyses reported in randomised trials: comparison of publications with protocols. *BMJ* 2008;**337**:a2299.

Chandler 2012

Chandler J, Churchill R, Higgins J, Lasserson T, Tovey D. Methodological standards for the reporting of new Cochrane Intervention Reviews Version 1.1, 17 December 2012. <http://www.editorial-unit.cochrane.org/mecir> 2012.

Clarke 2007

Clarke M. Standardising outcomes for clinical trials and systematic reviews. *Trials* 2007;**8**:39.

COMET Initiative

COMET Initiative. <http://www.comet-initiative.org/>.

Deeks 2011

Deeks JJ, Higgins JPT, Altman DG. Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

DerSimonian 1986

DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986;**7**(3):177–88.

Dickersin 1990

Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA* 1990;**263**(10):1385–9.

Dwan 2008

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**(8):e3081.

Dwan 2011

Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, Williamson PR. Comparison of protocols and registry entries to published reports for randomised controlled trials. *Cochrane Database of Systematic Reviews* 2011, Issue 1. [DOI: 10.1002/14651858.MR000031.pub2]

Dwan 2013b

Dwan K, Gamble C, Williamson PR, Kirkham JJ, Reporting Bias Group. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. *PLoS ONE* 2013;**8**(7):e66844.

Dwan 2014

Dwan K, Altman DG, Clarke M, Gamble C, Higgins JPT, Sterne JAC, et al. Evidence for the selective reporting of analyses and discrepancies in clinical trials: a systematic review of cohort studies of clinical trials. *PLoS Medicine* 2014;**11**(6):e1001666.

Freeman 1950

Freeman MF, Tukey JW. Transformations related to the angular and the square root. *Annals of Mathematical Statistics* 1950;**21**(4):607–11.

Hafdahl 2010a

Hafdahl AR. Article alerts: Introduction and items from 2009, Part I. *Research Synthesis Methods* 2010;**1**:81–7.

Hafdahl 2010b

Hafdahl AR. Article alerts: Items from 2009, Part II. *Research Synthesis Methods* 2010;**1**:319–26.

Hafdahl 2011a

Hafdahl AR. Article Alerts: items from 2010, Part II. *Research Synthesis Methods* 2011;**2**(4):279–86.

Hafdahl 2011b

Hafdahl AR. Article alerts: items from 2010, part I. *Research Synthesis Methods* 2011;**2**(2):131–8.

Hafdahl 2012

Hafdahl AR. Article Alerts: Items from 2011, Part I. *Research Synthesis Methods* 2012;**3**(4):325–31.

Higgins 2002

Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002;**21**(11):1539–58.

Higgins 2011a

Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Higgins 2011b

Higgins JPT, Altman DG, Sterne JAC. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Hutton 2000

Hutton JL, Williamson PR. Bias in meta-analysis due to outcome variable selection within studies. *Applied Statistics* 2000;**49**(3):359–70.

Kirkham 2010a

Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 2010; Vol. 340:e365.

Kirkham 2013

Kirkham JJ, Gargon E, Clarke M, Williamson PR. Can a core outcome set improve the quality of systematic reviews? -a survey of the Co-ordinating Editors of Cochrane Review Groups. *Trials* 2013;**14**:21.

Liberati 2009

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of Clinical Epidemiology* 2009;**62**(10):e1–34.

Mathieu 2009

Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 2009;**302**(9):977–84.

Moher 2009

Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Journal of Clinical Epidemiology* 2009;**62**(10):1006–12.

Moher 2012

Moher D, Stewart L, Shekelle P. Establishing a new journal for systematic review products. *Systematic Reviews* 2012;**1**:1.

Morissette 2011

Morissette K, Tricco AC, Horsley T, Chen MH, Moher D. Blinded versus unblinded assessments of risk of bias in studies included in a systematic review. *Cochrane Database of Systematic Reviews* 2011, Issue 9. [DOI: 10.1002/14651858.MR000025.pub2]

Norris 2013

Norris SL, Moher D, Reeves BC, Shea B, Loke Y, Garner S, et al. Issues relating to selective reporting when including non-randomized studies in systematic reviews on the effects of healthcare interventions. *Research Synthesis Methods* 2013;**4**(1):36–47.

Orsini 2006

Orsini N, Bottai M, Higgins J, Buchan I. HETEROGI: Stata module to quantify heterogeneity in a meta-analysis, Statistical Software Components. Boston College Department of Economics. <http://econpapers.repec.org/RePEc:boc:bocode:s449201>. 2006 (accessed 11 June 2014).

Oxman 1991

Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *Journal of Clinical Epidemiology* 1991;**44**(11):1271–8.

Page 2013a

Page MJ, McKenzie JE, Forbes A. Many scenarios exist for selective inclusion and reporting of results in randomized trials and systematic reviews. *Journal of Clinical Epidemiology* 2013;**66**(5):524–37.

PLoS 2011

PLoS Medicine Editors. Best practice in systematic reviews: the importance of protocols and registration. *PLoS Medicine* 2011;**8**(2):e1001009.

RevMan 2011 [Computer program]

The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.1. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2011.

Schünemann 2011

Schünemann HJ, Oxman AD, Vist GE, Higgins JPT, Deeks JJ, Glasziou P, et al. Chapter 12: Interpreting results and drawing conclusions. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Shamseer 2013

Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Protocols (PRISMA-P) 2013 (abstract). Oral presentation at the 21st Cochrane Colloquium; 2013 Sept 19-23; Quebec, Canada. 2013.

Shea 2007

Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology* 2007;**7**:10.

Song 2010

Song F, Parekh S, Hooper L, Loke YK, Ryder J, Sutton AJ, et al. Dissemination and publication of research findings: an updated review of related biases. *Health Technology Assessment* 2010;**14**(8):iii, ix–xi, 1–193.

SRC Methods Library 2013

US Agency for Healthcare Research and Quality Effective Healthcare Program's Scientific Resource Center Methods Library. <http://refworks.com/refshare2?site=040191157083200000/41331351619147490/SRC%20Methods%20-%20Production> (accessed 05 June 2013).

Sterne 2011

Sterne JAC, Egger M, Moher D. Chapter 10: Addressing reporting biases. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Stewart 2012

Stewart L, Moher D, Shekelle P. Why prospective registration of systematic reviews makes sense. *Systematic Reviews* 2012;**1**:7.

Tendal 2009

Tendal B, Higgins JPT, Jüni P, Hróbjartsson A, Trelle S, Nüesch E, et al. Disagreements in meta-analyses using outcomes measured on continuous or rating scales: observer agreement study. *BMJ* 2009;**339**:b3128.

Tricco 2008

Tricco AC, Tetzlaff J, Sampson M, Fergusson D, Cogo E, Horsley T, et al. Few systematic reviews exist documenting

- the extent of bias: a systematic review. *Journal of Clinical Epidemiology* 2008;**61**(5):422–34.
- Trikalinos 2013**
 Trikalinos TA, Trow P, Schmid CH. Simulation-Based Comparison of Methods for Meta-Analysis of Proportions and Rates [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US). <http://www.ncbi.nlm.nih.gov/books/NBK179162/>. 2013 (accessed 11 June 2014).
- Turner 2012**
 Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database of Systematic Reviews* 2012, Issue 11. [DOI: 10.1002/14651858.MR000030.pub2]
- Turner 2013**
 Turner L, Galipeau J, Garritty C, Manheimer E, Wieland LS, Yazdi F, et al. An evaluation of epidemiological and reporting characteristics of complementary and alternative medicine (CAM) systematic reviews (SRs). *PLoS ONE* 2013;**8**(1):e53536.
- Vedula 2009**
 Vedula SS, Bero L, Scherer RW, Dickersin K. Outcome reporting in industry-sponsored trials of gabapentin for off-label use. *New England Journal of Medicine* 2009;**361**(20):1963–71.
- Vedula 2013**
 Vedula SS, Li T, Dickersin K. Differences in reporting of analyses in internal company documents versus published trial reports: comparisons in industry-sponsored trials in off-label uses of gabapentin. *PLoS Medicine* 2013;**10**(1):e1001378.
- Viechtbauer 2010**
 Viechtbauer W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* 2010;**36**(3):1–48. <http://www.jstatsoft.org/v36/i03/paper>.
- Williamson 2005a**
 Williamson PR, Gamble C. Identification and impact of outcome selection bias in meta-analysis. *Statistics in Medicine* 2005;**24**(10):1547–61.
- Williamson 2005b**
 Williamson PR, Gamble C, Altman DG, Hutton JL. Outcome selection bias in meta-analysis. *Statistical Methods in Medical Research* 2005;**14**(5):515–24.
- Williamson 2012a**
 Williamson P, Altman D, Blazeby J, Clarke M, Gargon E. Driving up the quality and relevance of research through the use of agreed core outcomes. *Journal of Health Services Research and Policy* 2012;**17**(1):1–2.
- Williamson 2012b**
 Williamson PR, Clarke M. The COMET (Core Outcome Measures in Effectiveness Trials) Initiative: its role in improving Cochrane Reviews (editorial). *Cochrane Database of Systematic Reviews* 2012, (5). [DOI: 10.1002/14651858.ED000041]
- * Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Beller 2011

| | | |
|---------------------|--|-------------|
| Methods | Investigated discrepancies between systematic review abstracts and the main text; and selective partial reporting of outcomes in systematic review abstracts | |
| Data | <p>N = 182 (64 Cochrane and 118 non-Cochrane) systematic reviews</p> <p>Inclusion criteria:</p> <ol style="list-style-type: none"> 1. New comparative intervention reviews published Issue 4, 2009, of the Cochrane Database of Systematic Reviews or indexed in the National Library of Medicine's 119 Core Clinical Journals (Abridged Index Medicus) during 2009; 2. Included one or more meta-analyses comparing interventions. <p>Year(s) of publication of the systematic reviews: 2009</p> <p>Areas of health care addressed by the systematic reviews: Various (not specifically assessed)</p> <p>Methodological quality of systematic reviews: Not assessed</p> <p>Number (proportion) of systematic reviews that only included RCTs: Not assessed</p> <p>Number (proportion) of systematic reviews that are Cochrane Reviews: 64 (35%)</p> <p>Extent of overlap of RCTs included in the systematic reviews in the empirical study: not assessed, though unlikely to have impacted on the results</p> | |
| Comparisons | Abstract versus main text | |
| Outcomes | <p>Prevalence of systematic review abstracts with partially reported outcomes (i.e. size, direction, and precision not clearly reported)</p> <p>Association between partial reporting of outcomes and statistical significance of meta-analysis results</p> | |
| Notes | <p>Complete reporting defined as either the direction and size of effect can be determined from the wording or deduced numerically, and a measure of precision of the effect size is stated (e.g. CI) (minimum reporting standard), or the direction of effect is stated in words, the size of effect is stated in words and numerically and a measure of precision is stated (ideal reporting standard)</p> <p>Statistical significance defined as $P < 0.05$.</p> <p>Data on partial reporting were collected on the primary outcome of the systematic review or the first-reported outcome in the abstract if the primary outcome was not stated</p> <p>Same sample of Cochrane Review abstracts was included in Hopewell 2010.</p> <p>Additional outcomes measured but not included in the review were prevalence of abstracts reporting the magnitude, direction and statistical uncertainty (e.g. P value, 95% CI) of effect estimates (i.e. measures of 'incomplete reporting')</p> <p>Study published as a research letter in a journal and as a conference abstract. Additional unpublished data was retrieved from the authors</p> | |
| <i>Risk of bias</i> | | |
| Item | Authors' judgement | Description |

| | | |
|--|------------|---|
| <p>Is the empirical study at low risk of selection bias?</p> | <p>Yes</p> | <p>Quote: “The systematic reviews selected were all new reviews of interventions published in issue 4, 2009, of the Cochrane Library and from a search (based on Moher et al) of the National Library of Medicine’s 119 Core Clinical Journals (Abridged Index Medicus) during 2009. Reviews were eligible if they included 1 or more meta-analyses comparing interventions. Eligibility was broad, as there is no widely agreed definition of a systematic review. We used a definition by Moher et al: ” . . . the authors’ stated objective was to summarize evidence from multiple studies, and the article described explicit methods, regardless of the details provided.” One author (E.M. B.) screened all titles and abstracts with a second author (S.H.) independently checking citations classified as possible systematic reviews based on full-text review. All Cochrane Reviews and a random sample of eligible non-Cochrane reviews were selected for data extraction”</p> <p>Comment: This empirical study included all Cochrane Reviews registered during a specified time period, and a random sample of non-Cochrane systematic reviews</p> |
| <p>Is the empirical study at low risk of selective reporting bias?</p> | <p>Yes</p> | <p>Comment: Data for all outcomes and analyses specified in an unpublished study protocol (provided by the authors on request) were either reported in the publications or provided by the authors</p> |

Dwan 2013a

| | |
|----------------|---|
| <p>Methods</p> | <p>Investigated discrepancies between systematic review protocols and published systematic reviews</p> |
| <p>Data</p> | <p>N = 46 systematic reviews</p> <p>Inclusion criteria:</p> <ol style="list-style-type: none"> 1. Systematic reviews published by the Cochrane Cystic Fibrosis and Genetic Disorders (CFGD) group in the Cochrane Database of Systematic Reviews before 2010. <p>Exclusion criteria:</p> <ol style="list-style-type: none"> 1. Non-Cochrane systematic reviews or reviews published by a different Cochrane Review Group. <p>Year(s) of publication of the systematic reviews: 2006 to 2009</p> |

| | | |
|---|---|--|
| | <p>Areas of health care addressed by the systematic reviews: Cystic Fibrosis and genetic disorders</p> <p>Methodological quality of systematic reviews: Not assessed</p> <p>Number (proportion) of systematic reviews that only included RCTs: 7 reviews did not identify any RCTs, the remaining reviews only included RCTs</p> <p>Number (proportion) of systematic reviews that are Cochrane Reviews: 46 (100%)</p> <p>Extent of overlap of RCTs included in the systematic reviews in the empirical study: Not assessed, though unlikely to have impacted on the results</p> | |
| Comparisons | Systematic review protocols versus published systematic reviews | |
| Outcomes | <p>Prevalence of systematic reviews with:</p> <ul style="list-style-type: none"> any discrepancy in at least one outcome measure (primary, secondary or unlabelled) from protocol to full review; an upgrade of at least one outcome from secondary in the protocol to primary in the full review; a downgrade of at least one outcome from primary in the protocol to secondary in the full review; addition of at least one new outcome (primary, secondary, or unlabelled) in the review that was not specified at all in the protocol; omission of at least one outcome (primary, secondary, or unlabelled) in the review that was listed in the protocol. | |
| Notes | Study published as a journal article and conference abstract. Additional unpublished data retrieved from the authors | |
| Risk of bias | | |
| Item | Authors' judgement | Description |
| Is the empirical study at low risk of selection bias? | Yes | <p>Quote: "A cohort of systematic reviews published by the Cochrane Cystic Fibrosis and Genetic Disorders (CFGD) group on the Cochrane Library before 2010 were identified"</p> <p>Quote: "The CFGD group had 46 cystic fibrosis systematic reviews published as of 2010"</p> <p>Comment: This empirical study included all systematic reviews with an available protocol, which were published during a specified time period</p> |
| Is the empirical study at low risk of selective reporting bias? | Yes | <p>Comment: Data for all outcomes and analyses specified in an unpublished study protocol (provided by the authors on request) were either reported in the publications or provided by the authors</p> |

Hopewell 2010

| | |
|-------------|---|
| Methods | Investigated discrepancies between systematic review abstracts and the main text; and selective partial reporting of outcomes in systematic review abstracts |
| Data | <p>N = 62 systematic reviews</p> <p>Inclusion criteria:</p> <ol style="list-style-type: none"> 1. New reviews published in Issue 4, 2009 of the Cochrane Database of Systematic Reviews of Interventions where the primary outcome(s) were clearly stated in the full text and a meta-analysis had been conducted. <p>Exclusion criteria:</p> <ol style="list-style-type: none"> 1. Non-intervention reviews (reviews of methodology or diagnostic test accuracy) published in Issue 4, 2009 of the Cochrane Database of Systematic Reviews of Interventions. <p>Year(s) of publication of the systematic reviews: 2009</p> <p>Areas of health care addressed by the systematic reviews: Various (not specifically assessed)</p> <p>Methodological quality of systematic reviews: Not assessed</p> <p>Number (proportion) of systematic reviews that only included RCTs: Not assessed</p> <p>Number (proportion) of systematic reviews that are Cochrane Reviews: 62 (100%)</p> <p>Extent of overlap of RCTs included in the systematic reviews in the empirical study: Not assessed, though unlikely to have impacted on the results</p> |
| Comparisons | Abstract versus main text |
| Outcomes | <p>Prevalence of:</p> <ul style="list-style-type: none"> • systematic reviews with discrepancies in the number of primary and secondary outcomes reported between the abstract and main text; • systematic review abstracts where secondary outcomes were reported ahead of primary outcomes. <p>Association between reporting of outcomes and the statistical significance of the results</p> |
| Notes | <p>Statistical significance defined as $P < 0.05$.</p> <p>Same sample of systematic review abstracts was included in Beller 2011.</p> <p>Additional outcomes measured but not included in the review were proportion of abstracts reporting the magnitude, direction and statistical uncertainty (e.g. P value, 95% CI) of effect estimates (i.e. measures of 'incomplete reporting')</p> <p>Study published as a conference abstract only. Additional unpublished data retrieved from the authors</p> |

Risk of bias

| Item | Authors' judgement | Description |
|---|--------------------|---|
| Is the empirical study at low risk of selection bias? | Yes | <p>Quote: "We included all new reviews published in Issue 4, 2009 of The Cochrane Database of Systematic Reviews, where the primary outcome(s) were clearly stated in the full text and a meta-analysis had been conducted (n = 64); we excluded non-intervention reviews."</p> <p>Comment: This empirical study included</p> |

Hopewell 2010 (Continued)

| | | |
|---|---------|---|
| | | all systematic reviews published during a specified time period |
| Is the empirical study at low risk of selective reporting bias? | Unclear | Comment: No protocol for the study is available so it is unclear whether all measured and analysed outcomes were reported in the conference abstract and presented in the conference slides |

Kirkham 2010b

| | |
|-------------|--|
| Methods | Investigated discrepancies between systematic review protocols and published systematic reviews |
| Data | <p>N = 288 systematic reviews</p> <p>Inclusion criteria:</p> <ol style="list-style-type: none"> 1. New reviews published in the Cochrane Database of Systematic Reviews between Issue 4, 2006 and Issue 2, 2007, with an available protocol (either published or provided by the authors on request). <p>Exclusion criteria:</p> <ol style="list-style-type: none"> 1. Reviews from the Cochrane Methodology Review Group; 2. Reviews without an available protocol. <p>Year(s) of publication of the systematic reviews: 2006 to 2007</p> <p>Areas of health care addressed by the systematic reviews: Various (at least one review from each of 49 Cochrane Review Groups was included)</p> <p>Methodological quality of systematic reviews: Not assessed</p> <p>Number (proportion) of systematic reviews that only included RCTs: Not formally assessed (38 reviews did not identify any RCTs, and the remaining reviews included at least one RCT, though the proportion of reviews that included RCTs and non-RCTs is unclear)</p> <p>Number (proportion) of systematic reviews that are Cochrane Reviews: 288 (100%)</p> <p>Extent of overlap of RCTs included in the systematic reviews in the empirical study: Not assessed, though unlikely to have impacted on the results</p> |
| Comparisons | Systematic review protocols versus published systematic reviews |
| Outcomes | <p>Prevalence of systematic reviews with:</p> <ul style="list-style-type: none"> • any discrepancy in at least one outcome measure (primary, secondary, or unlabelled) from protocol to publication; • addition of at least one new primary outcome in the review that was not specified at all (i.e. as either a primary or secondary outcome) in the protocol; • omission of at least one primary outcome in the review that was listed as a primary outcome measure in the protocol; • an upgraded outcome (if a secondary protocol outcome was promoted to a review primary outcome; an upgrade also occurred if the review specified primary outcome(s) but the same outcomes listed in the protocol were not listed with any order of importance, i.e. primary or secondary); • a downgraded outcome (if a protocol primary outcome was demoted to a review secondary outcome; a downgrade also occurred if the protocol specified primary |

| | |
|--|---|
| | <p>outcome(s) but the same outcomes listed in the review had no order of importance). Risk of obtaining a statistically significant result ($P < 0.05$) in the meta-analysis if the discrepancy between protocol and published review was:</p> <ul style="list-style-type: none"> • either an “addition” or “upgrade” compared to if there was no discrepancy*; • a “downgrade” compared to if there was no discrepancy*. |
|--|---|

| | |
|-------|--|
| Notes | <p>Statistical significance defined as $P < 0.05$. *The analysis of this outcome is based on 245 meta-analysis results from 148 systematic reviews (140 of the 288 included systematic reviews were excluded from the analysis because they did not report a meta-analysis for the primary outcome). Rather than analysing one meta-analysis result per systematic review, the authors analysed all meta-analyses of the primary review comparison in each review (as many reviews had more than one primary outcome) Study published as a journal article and conference abstract. Additional unpublished data retrieved from the authors</p> |
|-------|--|

Risk of bias

| Item | Authors' judgement | Description |
|---|--------------------|--|
| Is the empirical study at low risk of selection bias? | Yes | Quote: “Eight percent (24/297) of reviews did not have a protocol sourced next to the review under the “Protocol and previous versions“ section on the Cochrane Library. The reason was not provided by two lead systematic reviewers. Seven (2% of 295 reviews) did not have a protocol: five reviewers went straight from registered title to review and two reviews were published by an alternative source and were later updated and developed into a Cochrane Review using Cochrane guidelines. For the remaining 15 reviews, the reviewer authors sent a copy of the protocol. These protocols were missing from the “Protocol and previous versions“ section of the Cochrane Library because a) the review was split into a number of separate reviews and only one protocol was registered (9 reviews), b) a draft protocol was accepted by the Cochrane Review Group (CRG) but was not registered on the Cochrane Library as it was never formally published (4 reviews) and c) the reviewer thought the protocol was registered on the Cochrane Library but its source location could not be found (2 reviews). For this last category, the CRG was contacted and the protocols had been withdrawn from the |

Kirkham 2010b (Continued)

| | | |
|---|-----|---|
| | | Library on the advice of the Collaboration because they were seen to be out of date. Thus 288 protocol-review pairs were available for study.” Comment: This empirical study included all systematic reviews with an available protocol, which were published during a specified time period |
| Is the empirical study at low risk of selective reporting bias? | Yes | Comment: Data for all outcomes and analyses specified in an unpublished study protocol (provided by the authors on request) were either reported in the publications or provided by the authors |

Parmelli 2007

| | |
|-------------|---|
| Methods | Investigated discrepancies between systematic review protocols and published systematic reviews |
| Data | N = 104 systematic reviews Inclusion criteria: 1. All new Cochrane systematic reviews published in Issues 2, 3, and 4, 2005, and Issue 1, 2006, of the Cochrane Database of Systematic Reviews with a protocol (methodology reviews eligible). Year(s) of publication of the systematic reviews: 2005 to 2006 Areas of health care addressed by the systematic reviews: Various (not specifically assessed) Methodological quality of systematic reviews: Not assessed Number (proportion) of systematic reviews that only included RCTs: Not assessed Number (proportion) of systematic reviews that are Cochrane Reviews: 104 (100%) Extent of overlap of RCTs included in the systematic reviews in the empirical study: Not assessed, though unlikely to have impacted on the results |
| Comparisons | Systematic review protocols versus published systematic reviews |
| Outcomes | Prevalence of systematic reviews with: <ul style="list-style-type: none"> • addition of at least one new outcome which was not defined in the protocol; • omission of at least one outcome which was defined in the protocol; • at least one “upgraded” outcome, i.e. outcomes defined as secondary in the protocol became primary in the systematic review or when not specified in the protocol became secondary or primary in the systematic review; • at least one “downgraded” outcome, i.e. outcomes stated as primary or secondary in the protocol became secondary or not specified in the systematic review, respectively. Odds of an outcome being statistically significant if the discrepancy between protocol and published review was: <ul style="list-style-type: none"> • an “upgrade” compared to if there was agreement between the protocol and full review*; • a “downgrade” compared to if there was agreement between the protocol and full |

Parmelli 2007 (Continued)

| | |
|-------|---|
| | <p>review*; <ul style="list-style-type: none"> • an “addition” compared to if there was agreement between the protocol and full review*. </p> |
| Notes | <p>Statistical significance defined as $P < 0.05$ *The analysis of this outcome is based on 336 meta-analyses from all 104 systematic reviews. The investigators reported that “only outcomes for which meta-analyses were carried out” were included in this analysis Additional outcomes measured but not included in the review were the proportion of outcomes that were discrepant between protocol and full review (i.e. discrepancies analysed at the outcome level, not the systematic review level) Study published as a conference abstract only. Additional unpublished data retrieved from the authors</p> |

Risk of bias

| Item | Authors' judgement | Description |
|---|--------------------|--|
| Is the empirical study at low risk of selection bias? | Yes | <p>Quote: “All new SRs published in the 2005 CL Issues 2, 3, 4 and in 2006 CL Issue 1 were identified. A random sample of SRs was selected.” Comment: This empirical study included a random sample of systematic reviews</p> |
| Is the empirical study at low risk of selective reporting bias? | Unclear | <p>Comment: No protocol for the study is available so it is unclear whether all measured and analysed outcomes were reported in the conference abstract and presented in the conference slides</p> |

Silagy 2002

| | |
|---------|--|
| Methods | Investigated discrepancies between systematic review protocols and published systematic reviews |
| Data | <p>N = 47 systematic reviews Inclusion criteria: 1. All new Cochrane Reviews appearing for the first time in the Cochrane Database of Systematic Reviews, Issue 3, 2000; 2. Protocol published for the review in the previous issue of the Cochrane Database of Systematic Reviews, Issue 2, 2000. Exclusion criteria: 1. Updated Cochrane Reviews appearing in the Cochrane Database of Systematic Reviews, Issue 3, 2000. Year(s) of publication of the systematic reviews: 2000 Areas of health care addressed by the systematic reviews: Various (not specifically assessed) Methodological quality of systematic reviews: Not assessed Number (proportion) of systematic reviews that only included RCTs: Not assessed</p> |

Silagy 2002 (Continued)

| | | |
|---|--|---|
| | Number (proportion) of systematic reviews that are Cochrane Reviews: 47 (100%) Extent of overlap of RCTs included in the systematic reviews in the empirical study: Not assessed, though unlikely to have impacted on the results | |
| Comparisons | Systematic review protocols versus published systematic reviews | |
| Outcomes | Prevalence of systematic reviews which: <ul style="list-style-type: none"> • added new outcomes from the protocol to the published review; • added new measurement methods for existing outcomes from the protocol to the published review; • omitted outcomes from the protocol to the published review; • made changes to any of the different sections (i.e. background, objectives, types of studies, types of outcomes, etc.) of the protocols and reviews based on the results (as reported by authors when contacted to provide reasons for changes). | |
| Notes | Additional outcomes measured but not included in the review were the prevalence of systematic reviews with discrepancies in sections other than outcomes and analyses (i.e. background section, search strategy) Study published as a journal article and conference abstract. No additional data retrieved from authors | |
| <i>Risk of bias</i> | | |
| Item | Authors' judgement | Description |
| Is the empirical study at low risk of selection bias? | Yes | Quote: "We identified 66 completed reviews appearing for the first time in the Cochrane Library; 2000, issue 3. Of these, we found a protocol for 71% (n = 47) in 2000, issue 2, of the Cochrane Library. None of these published protocols had any external comments or criticisms attached. The remaining 19 reviews had apparently been carried out without prior publication of a protocol." Comment: This empirical study included all Cochrane Reviews published in a single issue of <i>The Cochrane Library</i> that had a published protocol. |
| Is the empirical study at low risk of selective reporting bias? | Unclear | Comment: No protocol for the study is available so it is unclear whether all measured and analysed outcomes were reported in the publication |

Vlassov 2008

| | | |
|---|--|--|
| Methods | Investigated discrepancies between systematic review abstracts and the main text | |
| Data | <p>N = 100 systematic reviews</p> <p>Inclusion criteria:</p> <ol style="list-style-type: none"> 1. New, updated, and existing systematic reviews published in Issue 1, 2008 of the Cochrane Database of Systematic Reviews with included studies. <p>Exclusion criteria:</p> <ol style="list-style-type: none"> 1. Cochrane Methodology Reviews; 2. Cochrane Reviews without included studies. <p>Year(s) of publication of the systematic reviews: 1996 to 2007</p> <p>Areas of health care addressed by the systematic reviews: Various (not specifically assessed)</p> <p>Methodological quality of systematic reviews: Not assessed</p> <p>Number (proportion) of systematic reviews that only included RCTs: Not assessed</p> <p>Number (proportion) of systematic reviews that are Cochrane Reviews: 100 (100%)</p> <p>Extent of overlap of RCTs included in the systematic reviews in the empirical study: Not assessed, though unlikely to have impacted on the results</p> | |
| Comparisons | Abstract versus main text | |
| Outcomes | <p>Prevalence of systematic review abstracts which:</p> <ul style="list-style-type: none"> • report the results of other outcomes before the results of the primary outcome; • report the results of other outcomes before the results of the primary outcome (where the other outcome was statistically significant); • report a primary outcome differently to the primary outcome reported in the main text (defined as different phrasing and unclear description). | |
| Notes | <p>Statistical significance defined as $P < 0.05$.</p> <p>Additional outcomes measured but not included in the review were prevalence of abstracts reporting the magnitude, direction and statistical uncertainty (e.g. P value, 95% CI) of effect estimates (i.e. measures of 'incomplete reporting')</p> <p>Study published as a conference abstract only. Additional unpublished data retrieved from the author</p> | |
| <i>Risk of bias</i> | | |
| Item | Authors' judgement | Description |
| Is the empirical study at low risk of selection bias? | Yes | <p>Quote: "It was not a random, but really 'regular' sample: having the DB [database] open in alphabet order I selected first 10 of every 100 abstracts. Reviews without trials found were excluded" (personal communication)</p> <p>Comment: This empirical study included a representative systematic sample of systematic reviews</p> |
| Is the empirical study at low risk of selective reporting bias? | Unclear | <p>Comment: No protocol for the study is available so it is unclear whether all mea-</p> |

| | | |
|--|--|--|
| | | sured and analysed outcomes were reported in the conference abstract |
|--|--|--|

Characteristics of excluded studies [ordered by study ID]

| Study | Reason for exclusion |
|-----------------|---|
| Assendelft 1995 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Auperin 1997 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Aytug 2012 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Bhandari 2001 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Bjordal 2003 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Bjordal 2005 | Investigated the extent of multiplicity of data in trial reports and the impact of multiplicity on meta-analysis results in one systematic review, not whether results were selectively included by the review authors |
| Bohlius 2005 | Investigated the extent of multiplicity of data in trial reports and the impact of multiplicity on meta-analysis results in one systematic review, not whether results were selectively included by the review authors |
| Bow 2010 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| CEU 2011 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| CEU 2012 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Choi 2001 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| da Costa 2013 | Investigated the extent of multiplicity of final and change from baseline data in trial reports and the potential impact of inclusion of the most positive versus most negative value in meta-analyses, but not whether results were selectively included by the review authors |

(Continued)

| | |
|-------------------|--|
| Delaney 2005 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Dundar 2006 | Investigated the extent of multiplicity of data in trial reports and the impact of multiplicity on meta-analysis results in one systematic review, not whether results were selectively included by the review authors |
| Faggion Jr 2014 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Farmer 2012 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Fishbain 2000 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Gianola 2013 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Hartling 2004 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Hopewell 2008 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Jørgensen 2006 | Investigated discrepancies in meta-analytic effect estimates between pairs of systematic reviews of the same topic, but the relative extent to which the discrepancies were caused by selective inclusion or RCT data versus differential inclusion of trials resulting from slightly different eligibility criteria or search strategy was not explored |
| Kelly 2001 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Kuukasjärvi 2006a | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Kuukasjärvi 2006b | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Lacasse 1999 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Ma 2012 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Minozzi 2006a | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |

(Continued)

| | |
|----------------|--|
| Minozzi 2006b | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Moher 2007 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Moseley 2009 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Roundtree 2009 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Sacks 1987 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Sacks 1992 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Sacks 1996 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Schwarzer 2001 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Shea 2006 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Sheikh 2007 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Song 1997 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Stroup 2001 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |
| Tendal 2011 | Investigated the extent of multiplicity of data in trial reports and the impact of multiplicity on meta-analysis results in 19 systematic reviews, not whether results were selectively included by the systematic reviewers |
| Wee 2008 | Investigated methods or reporting of systematic reviews using a quality checklist that did not include item(s) focusing on selective inclusion or reporting |

21 more studies were excluded, see main text

Characteristics of studies awaiting assessment *[ordered by study ID]*

Johnston 2012

| | |
|-------------|---|
| Methods | Assessed in Cochrane and non-Cochrane systematic reviews reporting dichotomous outcomes: (a) the concordance between abstracts and full texts in the reporting of patient important outcomes, and (b) the extent to which absolute estimates of effect are reported in the abstract |
| Data | 200 systematic reviews (Cochrane and non-Cochrane) published in 2010 |
| Comparisons | Outcomes reported in the full text compared to outcomes reported in the summary of systematic reviews |
| Outcomes | Data abstraction ongoing at the time of submission of abstract to 20th Cochrane Colloquium, 2012 |
| Notes | |

Middleton 2010

| | |
|-------------|---|
| Methods | Assessed all relevant Cochrane Reviews for their ability to identify interventions with the potential to reduce stillbirths or to reduce factors known to be associated with stillbirths in high income countries |
| Data | 254 systematic reviews published in the Cochrane Database of Systematic Reviews, Issue 4, 2009 |
| Comparisons | Unclear |
| Outcomes | Only reported that “There was some evidence of selective outcome reporting bias at both the review and trial level” |
| Notes | |

Characteristics of ongoing studies *[ordered by study ID]*

Page 2013b

| | |
|---------------------|---|
| Trial name or title | An empirical investigation of the potential impact of selective inclusion of results in systematic reviews of interventions |
| Methods | The primary objectives of this study are to investigate in a cohort of systematic reviews of randomised controlled trials of interventions for rheumatoid arthritis, osteoarthritis, depressive disorders and anxiety disorders: (i) how often there is multiplicity of outcome data in trial reports; (ii) the association between selection of trial outcome data included in a meta-analysis and the magnitude and statistical significance of the trial result, and; (iii) the impact of the selection of outcome data on meta-analytic results |
| Data | - |
| Comparisons | Selective inclusion of RCT outcome data in systematic reviews |
| Outcomes | - |

Page 2013b (Continued)

| | |
|---------------------|---|
| Starting date | 10 April 2013 |
| Contact information | Matthew Page, Monash University (matthew.page@monash.edu) |
| Notes | |

DATA AND ANALYSES

Comparison 1. Association between outcome discrepancies and statistical significance

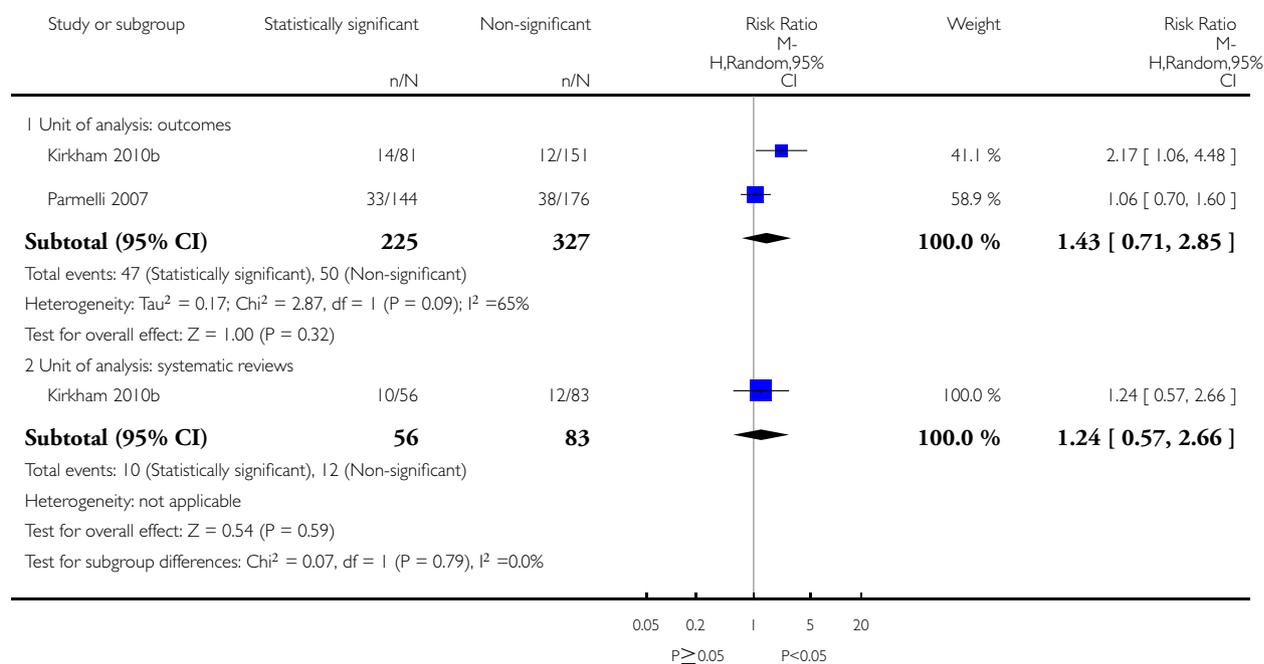
| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|--|----------------|---------------------|----------------------------------|-------------------|
| 1 Outcome "added/upgraded" | 2 | | Risk Ratio (M-H, Random, 95% CI) | Subtotals only |
| 1.1 Unit of analysis: outcomes | 2 | 552 | Risk Ratio (M-H, Random, 95% CI) | 1.43 [0.71, 2.85] |
| 1.2 Unit of analysis: systematic reviews | 1 | 139 | Risk Ratio (M-H, Random, 95% CI) | 1.24 [0.57, 2.66] |
| 2 Outcome "downgraded" | 2 | | Risk Ratio (M-H, Random, 95% CI) | Subtotals only |
| 2.1 Unit of analysis: outcomes | 2 | 484 | Risk Ratio (M-H, Random, 95% CI) | 1.26 [0.60, 2.62] |
| 2.2 Unit of analysis: systematic reviews | 1 | 126 | Risk Ratio (M-H, Random, 95% CI) | 0.79 [0.21, 3.00] |

Analysis 1.1. Comparison 1 Association between outcome discrepancies and statistical significance, Outcome 1 Outcome "added/upgraded".

Review: Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions

Comparison: 1 Association between outcome discrepancies and statistical significance

Outcome: 1 Outcome "added/upgraded"

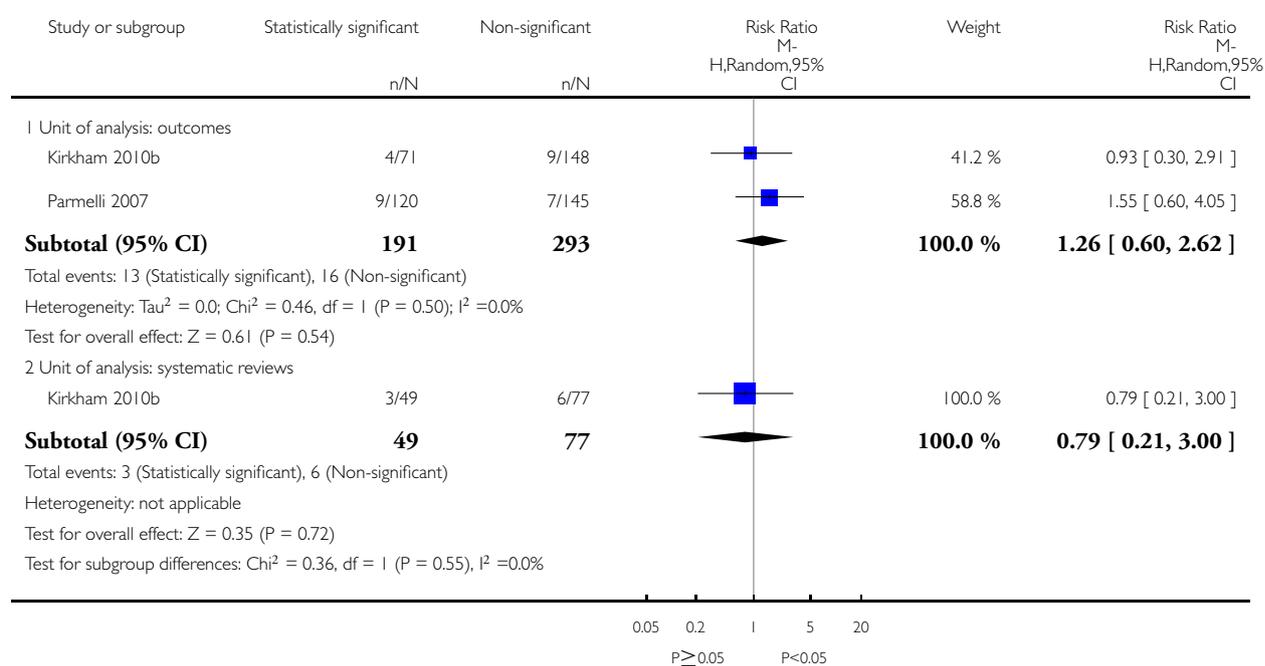


Analysis 1.2. Comparison 1 Association between outcome discrepancies and statistical significance, Outcome 2 Outcome "downgraded".

Review: Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions

Comparison: 1 Association between outcome discrepancies and statistical significance

Outcome: 2 Outcome "downgraded"



APPENDICES

Appendix I. Possible specific estimates of prevalence of selective inclusion and reporting in systematic reviews

Empirical studies investigating selective inclusion of RCT outcome data in systematic reviews

For each empirical study investigating selective inclusion of RCT outcome data in systematic reviews, we will report, depending on the data available, the prevalence of systematic reviews in which the review authors (as determined by the empirical study investigators) have selectively:

- omitted all data for a measured outcome;
- included specific data for an outcome (e.g. inclusion of a subset of outcome data measured using multiple measurement scales, or at multiple time points);
- included specific analyses of the same outcome data (e.g. inclusion of data analysed based on per-protocol versus intention-to-treat analysis);
- included subsets of outcome data (e.g. inclusion of subgroup analyses or sub-scale data) reported in the RCT publications.

Empirical studies comparing outcomes or analyses reported in systematic review registry entries to outcomes or analyses reported in systematic reviews

For each empirical study comparing systematic review registry entries to publications, we will report, depending on the data available, the prevalence of discrepancies as follows:

- a primary outcome stated in the registry entry is downgraded to secondary in the publication;
- a secondary outcome in the registry entry is upgraded to primary in the publication;
- a new primary outcome that was not stated in the registry entry (as primary or secondary) is added to the publication;
- a new secondary outcome that was not stated in the registry entry (as primary or secondary) is added to the publication;
- a primary outcome stated in the registry entry is omitted from the publication;
- a secondary outcome stated in the registry entry is omitted from the publication;
- the definition of a primary outcome is different (although being the same variable) in the registry entry compared to the publication;
- the definition of a secondary outcome is different (although being the same variable) in the registry entry compared to the publication;
- a primary outcome stated in the registry entry is unspecified as primary or secondary in the publication;
- a secondary outcome stated in the registry entry is unspecified as primary or secondary in the publication;
- the outcome used in the power calculation differs from the registry entry to the publication;
- the time points that primary outcomes are reported as being measured at differ from the registry entry to the publication;
- the time points that secondary outcomes are reported as being measured at differ from the registry entry to the publication;
- an analysis (e.g. method to handle missing data or confounders to include in adjusted analyses) stated in the registry entry is omitted from the publication;
- an analysis (e.g. method to handle missing data or confounders to include in adjusted analyses) not stated in the registry entry is included in the publication.

Empirical studies comparing outcomes or analyses reported in systematic review protocols to outcomes or analyses reported in systematic reviews

For each empirical study comparing systematic review protocols to publications, we will report, depending on the data available, the prevalence of discrepancies as for the registry entries (see just above).

Empirical studies comparing outcomes or analyses reported in the full text to outcomes reported in the summary within systematic reviews

For each empirical study comparing the full text of the systematic review with at least one summary (i.e. abstract, plain language summary or 'Summary of findings' table) of the systematic review, we will report, depending on the data available, the prevalence of systematic reviews in which the review authors have:

- omitted from the summary a primary outcome reported in the full text;
- upgraded an outcome reported in the full text as secondary to primary;
- downgraded an outcome reported in the full text as primary to secondary;
- reported the results of primary or secondary outcomes inconsistently across the full text and summary;
- reported the results of a secondary outcome before the results of the primary outcome(s) in the summary.

Empirical studies investigating selective partial reporting of outcomes in the full text or summary of systematic reviews

For each empirical study investigating the selective partial reporting of outcomes in the full text or summary (i.e. abstract, plain language summary or 'Summary of findings' table) of the systematic review, we will report, depending on the data available, the prevalence of:

- systematic reviews in which the review authors have partially reported at least one primary outcome based on the results;
- systematic reviews in which the review authors have partially reported at least one secondary outcome based on the results.

Appendix 2. Search strategies

OvidSP MEDLINE

1. Clinical Protocols/
2. protocol\$.ti,ab.
3. regist\$.ti,ab.
4. Registries/
5. or/1-4
6. Meta-Analysis as Topic/
7. meta-analy\$.ti,ab.
8. meta analy\$.ti,ab.
9. metaanaly\$.ti,ab.
10. (systematic\$ adj (review\$ or overview\$)).ti,ab.
11. (quantitativ\$ adj (review\$ or overview\$)).ti,ab.
12. meta-synthes\$.ti,ab.
13. meta synthes\$.ti,ab.
14. metasyntes\$.ti,ab.
15. mega-synthes\$.ti,ab.
16. mega synthes\$.ti,ab.
17. megasyntes\$.ti,ab.
18. or/6-17
19. "Bias (Epidemiology)"/
20. publication bias/
21. (unreported or "incompletely reported" or "partially reported" or "fully reported" or "not reported" or "non-report\$" or missing or omission or omit\$ or "not publish\$").ti,ab.
22. ((selectiv\$ or suppress\$ or non\$ or bias\$) adj5 (report\$ or publish\$ or publication\$ or inclu\$)).ti,ab.
23. or/19-22
24. (discrepan\$ adj5 (protocol\$ or regist\$)).ti,ab.
25. (compar\$ adj8 publication\$ adj8 protocol\$).ti,ab.
26. (compar\$ adj8 protocol\$ adj8 publication\$).ti,ab.
27. (publication\$ adj8 protocol\$ adj8 compar\$).ti,ab.
28. (publication\$ adj8 compar\$ adj8 protocol\$).ti,ab.
29. (protocol\$ adj8 publication\$ adj8 compar\$).ti,ab.

30. (protocol\$ adj8 compar\$ adj8 publication\$).ti,ab.
31. (compar\$ adj8 publication\$ adj8 regist\$).ti,ab.
32. (compar\$ adj8 regist\$ adj8 publication\$).ti,ab.
33. (publication\$ adj8 regist\$ adj8 compar\$).ti,ab.
34. (publication\$ adj8 compar\$ adj8 regist\$).ti,ab.
35. (regist\$ adj8 publication\$ adj8 compar\$).ti,ab.
36. (regist\$ adj8 compar\$ adj8 publication\$).ti,ab.
37. or/24-36
38. 5 and 18 and 23
39. 18 and 37
40. 38 or 39
41. cochrane database of systematic reviews.jn.
42. 40 not 41

OvidSP EMBASE

1. Clinical Protocols/
2. protocol\$.ti,ab.
3. regist\$.ti,ab.
4. Register/
5. or/1-4
6. Meta Analysis/
7. meta analy\$.ti,ab. or meta-analy\$.ti,ab. or metaanaly\$.ti,ab.
8. systematic\$ adj (review\$ or overview\$).ti,ab.
9. quantitativ\$ adj (review\$ or overview\$).ti,ab.
10. meta synthes\$.ti,ab. or meta-synthes\$.ti,ab. or metasynthes\$.ti,ab.
11. mega synthes\$.ti,ab. or mega-synthes\$.ti,ab. or megasynthes\$.ti,ab.
12. or/6-11
13. publishing/
14. (unreported or “incompletely reported” or “partially reported” or “fully reported” or “not reported” or “non-report\$” or missing or omission or omit\$ or “not publish\$”).ti,ab.
15. ((selectiv\$ or suppress\$ or non\$ or bias\$) adj5 (report\$ or publish\$ or publication\$ or inclu\$)).ti,ab.
16. or/13-15
17. (discrepan\$ adj5 (protocol\$ or regist\$)).ti,ab.
18. (compar\$ adj8 publication\$ adj8 protocol\$).ti,ab.
19. (compar\$ adj8 protocol\$ adj8 publication\$).ti,ab.
20. (publication\$ adj8 protocol\$ adj8 compar\$).ti,ab.
21. (publication\$ adj8 compar\$ adj8 protocol\$).ti,ab.
22. (protocol\$ adj8 publication\$ adj8 compar\$).ti,ab.
23. (protocol\$ adj8 compar\$ adj8 publication\$).ti,ab.
24. (compar\$ adj8 publication\$ adj8 regist\$).ti,ab.
25. (compar\$ adj8 regist\$ adj8 publication\$).ti,ab.
26. (publication\$ adj8 regist\$ adj8 compar\$).ti,ab.
27. (publication\$ adj8 compar\$ adj8 regist\$).ti,ab.
28. (regist\$ adj8 publication\$ adj8 compar\$).ti,ab.
29. (regist\$ adj8 compar\$ adj8 publication\$).ti,ab.
30. or/17-29
31. 5 and 12 and 16
32. 12 and 30
33. 31 or 32
34. cochrane database of systematic reviews.jn.
35. 33 not 34

OvidSP PsycINFO

1. protocol\$.ti,ab.
2. regist\$.ti,ab.
3. 1 or 2
4. Meta Analysis/
5. (meta analy\$ or meta-analy\$ or metaanaly\$).ti,ab.
6. (systematic\$ adj (review\$ or overview\$)).ti,ab.
7. (quantitativ\$ adj (review\$ or overview\$)).ti,ab.
8. (meta synthes\$ or meta-synthes\$ or metasynthes\$).ti,ab.
9. (mega synthes\$ or mega-synthes\$ or megasynthes\$).ti,ab.
10. or/4-9
11. (unreported or “incompletely reported” or “partially reported” or “fully reported” or “not reported” or “non-report\$” or missing or omission or omit\$ or “not publish\$”).ti,ab.
12. ((selectiv\$ or suppress\$ or non\$ or bias\$) adj5 (report\$ or publish\$ or publication\$ or inclu\$)).ti,ab.
13. 11 or 12
14. (discrepan\$ adj5 (protocol\$ or regist\$)).ti,ab.
15. (compar\$ adj8 publication\$ adj8 protocol\$).ti,ab.
16. (compar\$ adj8 protocol\$ adj8 publication\$).ti,ab.
17. (publication\$ adj8 protocol\$ adj8 compar\$).ti,ab.
18. (publication\$ adj8 compar\$ adj8 protocol\$).ti,ab.
19. (protocol\$ adj8 publication\$ adj8 compar\$).ti,ab.
20. (protocol\$ adj8 compar\$ adj8 publication\$).ti,ab.
21. (compar\$ adj8 publication\$ adj8 regist\$).ti,ab.
22. (compar\$ adj8 regist\$ adj8 publication\$).ti,ab.
23. (publication\$ adj8 regist\$ adj8 compar\$).ti,ab.
24. (publication\$ adj8 compar\$ adj8 regist\$).ti,ab.
25. (regist\$ adj8 publication\$ adj8 compar\$).ti,ab.
26. (regist\$ adj8 compar\$ adj8 publication\$).ti,ab.
27. or/14-26
28. 3 and 10 and 13
29. 10 and 27
30. 28 or 29

Cochrane Methodology Register (Wiley InterScience (Online))

1. (protocol* OR regist*):ti in Methods Studies
2. (protocol* OR regist*):ab in Methods Studies
3. (#1 OR #2)
4. (meta-analy* OR meta analy* OR metaanaly* OR meta-synthes* OR meta synthes* OR metasynthes* OR mega-synthes* OR mega synthes* OR megasynthes*):ti in Methods Studies
5. (meta-analy* OR meta analy* OR metaanaly* OR meta-synthes* OR meta synthes* OR metasynthes* OR mega-synthes* OR mega synthes* OR megasynthes*):ab in Methods Studies
6. systematic* adj (review* or overview*):ti in Methods Studies
7. systematic* adj (review* or overview*):ab in Methods Studies
8. quantitativ* adj (review* or overview*):ti in Methods Studies
9. quantitativ* adj (review* or overview*):ab in Methods Studies
10. (#4 OR #5 OR #6 OR #7 OR #8 OR #9)
11. “bias in meta-analysis”:kw in Methods Studies
12. (“study identification” next “publication bias”):kw in Methods Studies
13. (unreported OR “incompletely reported” OR “partially reported” OR “fully reported” OR “not reported” OR “non reported” OR “non-reported” OR “non reporting” OR “nonreporting” OR missing OR omission OR “not published” OR “not publishing” OR “not included” OR “non included” OR “non-included” OR “not including” OR “non inclusion” OR “selectively included” OR “selective inclusion” OR “partially included” OR “partial inclusion”):ti in Methods Studies

14. (unreported OR “incompletely reported” OR “partially reported” OR “fully reported” OR “not reported” OR “non reported” OR “non-reported” OR “non reporting” OR “nonreporting” OR missing OR omission OR “not published” OR “not publishing” OR “not included” OR “non included” OR “non-included” OR “not including” OR “non inclusion” OR “selectively included” OR “selective inclusion” OR “partially included” OR “partial inclusion”):ab in Methods Studies
15. omit*:ti in Methods Studies
16. omit*:ab in Methods Studies
17. ((selectiv* OR suppress* OR non* OR bias*) NEAR/5 (report* OR publish* OR publication* OR inclu*)):ti in Methods Studies
18. ((selectiv* OR suppress* OR non* OR bias*) NEAR/5 (report* OR publish* OR publication* OR inclu*)):ab in Methods Studies
19. (#11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18)
20. (discrepan* NEAR/5 (protocol* OR regist*)):ti in Methods Studies
21. (discrepan* NEAR/5 (protocol* OR regist*)):ab in Methods Studies
22. (compar* NEAR/8 publication* NEAR/8 protocol*):ti in Methods Studies
23. (compar* NEAR/8 publication* NEAR/8 protocol*):ab in Methods Studies
24. (compar* NEAR/8 publication* NEAR/8 regist*):ti in Methods Studies
25. (compar* NEAR/8 publication* NEAR/8 regist*):ab in Methods Studies
26. (#20 OR #21 OR #22 OR #23 OR #24 OR #25)
27. (#3 AND #10 AND #19)
28. (#10 AND #26)
29. (#19 OR #27 OR #28)

Web of Science

1. TS=protocol*
2. TS=registr*
3. #1 OR #2
4. TS=meta analy*
5. TS=meta-analy*
6. TS=metaanaly*
7. TS=“systematic* review*”
8. TS=“systematic* overview*”
9. TS=“quantitativ* review*”
10. TS=“quantitativ* overview*”
11. TS=meta synthes*
12. TS=meta-synthes*
13. TS=metasynthes*
14. TS=mega synthes*
15. TS=mega-synthes*
16. TS=megasynthes*
17. #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16
18. TS=“Bias (Epidemiology)”
19. TS=publication bias
20. TS=(unreported or “incompletely reported” or “partially reported” or “fully reported” or “not reported” or “non-report*” or missing or omission or omit* or “not publish*”)
21. TS=((selective* or suppress* or non* or bias*) NEAR/5 (report* or publish* or publication* or inclu*))
22. #18 OR #19 OR #20 OR #21
23. TS=(discrepan* SAME (protocol* or registr*))
24. TS=(compare* SAME publication* SAME protocol*)
25. TS=(compare* SAME protocol* SAME publication*)
26. TS=(publication* SAME protocol* SAME compar*)
27. TS=(publication* SAME compar* SAME protocol*)
28. TS=(protocol* SAME publication* SAME compar*)
29. TS=(protocol* SAME compar* SAME publication*)

30. TS=(compar* SAME publication* SAME registr*)
31. TS=(compare* SAME registr* SAME publication*)
32. TS=(publication* SAME registr* SAME compar*)
33. TS=(publication* SAME compar* SAME registr*)
34. TS=(registr* SAME publication* SAME compar*)
35. TS=(registr* SAME compar* SAME publication*)
36. #23 OR #24 OR 25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35
37. #3 AND #17 AND #22
38. #17 AND #36
39. #37 OR #38

F E E D B A C K

Eligibility criteria for the review, 15 October 2014

Summary

Feedback submitted by Professor Lisa Bero:

I am wondering why the following study did not appear among the included or excluded studies in your review. It examines whether drug trial outcomes that were submitted to the FDA but not published (ie, selectively reported outcomes) change the effects of meta-analyses if they are then included in the meta-analyses: Hart, B, Lundh, A and Bero, L. The effect of reporting bias on meta-analyses of drug trials: Re-analysis of meta-analyses. *BMJ*, 2011;343:d7202. doi: 10.1136/bmj.d7202

Since unpublished drug trial data are readily available from drug regulatory agencies, failure to look for such data can be considered a type of selection of data. I suggest that the update of your review expand the scope to assess the impact of not including unpublished data that are available from drug regulatory authorities.

Reply

The study by Hart and colleagues investigates whether selective reporting in trials affects the results of meta-analyses if the unreported data can be identified and then included in the relevant meta-analyses. Hart and colleagues identified the unpublished data and found that the results changed when they re-analysed the meta-analyses. However, it is unclear whether the authors of the meta-analyses were aware of the unpublished drug trial data and deliberately excluded it from the meta-analyses, or were even aware that unpublished drug trial data is readily available from drug regulatory agencies. Therefore, we decided that the study by Hart and colleagues did not fit the scope of our review.

We concur with Professor Bero that failure to look for unpublished drug trial data that is readily available could be considered a form of selective exclusion of outcome data. When the review is updated, we will consider expanding the scope to include studies assessing the impact of not including unpublished data that are available from drug regulatory agencies.

Contributors

Matthew J. Page

WHAT'S NEW

Last assessed as up-to-date: 8 May 2013.

| Date | Event | Description |
|-----------------|--------------------------------|---|
| 30 January 2015 | Feedback has been incorporated | Addition of feedback and response about the eligibility criteria for the review |

CONTRIBUTIONS OF AUTHORS

MJP and JEM developed the protocol, with comments from JK, KD, SG and AF. MJP, JEM, JK, KD and SK screened search results to identify eligible studies. MJP, JEM, KD, JK and SK extracted data from and assessed risk of bias of included studies. MJP and JEM carried out the analysis. MJP drafted and all authors commented on the review.

DECLARATIONS OF INTEREST

JK and KD are lead authors of two empirical studies that we have included in the review ([Dwan 2013a](#); [Kirkham 2010b](#)) but were not involved in the data extraction or risk of bias assessment. MJP, JEM, SG and AF are authors of an ongoing study of selective inclusion of results in systematic reviews ([Page 2013b](#)).

SOURCES OF SUPPORT

Internal sources

- Australasian Cochrane Centre, School of Public Health & Preventive Medicine, Monash University, Australia.

External sources

- Australian Postgraduate Award, Australia.

This work was conducted as part of a PhD undertaken by MJP, which is funded by an Australian Postgraduate Award administered through Monash University, Australia

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

In the [Types of methods](#) section of our protocol, we described five types of selective inclusion and reporting of outcomes and analyses that we planned to focus on and described the fifth as “partial reporting in systematic reviews”. We have re-labelled this as “selective partial reporting in systematic reviews” as we were only interested in studies investigating partial reporting of outcomes that is based on their results (e.g. whether statistically significant outcomes are more likely to be completely reported than non-significant outcomes). This is in contrast to studies investigating the “reporting quality” of systematic reviews by determining how often outcomes are partially reported (e.g. the proportion of systematic reviews not reporting an effect estimate or CI for outcomes) regardless of the results of those outcomes, which is outside the scope of our review.

The search strategies reported in the protocol were all modified. The original Cochrane Methodology Register search strategy included a step that required “systematic reviews or meta-analysis” terms to be combined with “selective inclusion and reporting bias” terms (using the 'AND' Boolean operator). Inclusion of this step made the Cochrane Methodology Register search strategy insensitive, as studies we were already aware of ([Dwan 2013a](#); [Parmelli 2007](#); [Silagy 2002](#); [Vlassov 2008](#)) were not identified. Therefore, we modified the Cochrane Methodology Register search strategy by allowing the “selective inclusion and reporting bias” terms to be searched alone

(i.e. they were not combined with “systematic review or meta-analysis” terms). In contrast, the step that required “systematic reviews or meta-analysis” terms to be combined with “selective inclusion and reporting bias” terms was removed from each of the Ovid MEDLINE, Ovid EMBASE, Ovid PsycInfo and ISI Web of Science search strategies because it was overly sensitive. Inclusion of this step in the searches of these four databases yielded a very large number of ineligible citations of systematic reviews of intervention studies that reported terms such as ‘bias’ or ‘selective reporting’ in the abstract. Therefore, the original step in these four databases was replaced with a step combining “systematic reviews or meta-analysis” terms with “selective inclusion and reporting bias” terms and “protocol or register” terms, which is similar to the search strategy used in [Dwan 2011](#). We reported our updated search strategies in [Appendix 2](#). We also searched the US AHRQ Effective Healthcare Program’s SRC Methods Library ([SRC Methods Library 2013](#)); abstract books of the 2011 and 2012 Cochrane Colloquia (which were not indexed in any electronic databases at the time of the search); and article alerts for methodological work in research synthesis published from 2009 to 2011, which are available in *Research Synthesis Methods* ([Hafdahl 2010b](#); [Hafdahl 2010a](#); [Hafdahl 2011a](#); [Hafdahl 2011b](#); [Hafdahl 2012](#)).

INDEX TERMS

Medical Subject Headings (MeSH)

*Meta-Analysis as Topic; *Randomized Controlled Trials as Topic; *Review Literature as Topic; *Selection Bias; Treatment Outcome