Research Techniques Made Simple: Assessing Risk of Bias in Systematic Reviews

Aaron M. Drucker[1], Patrick Fleming[2] and An-Wen Chan[3]

Systematic reviews are increasingly utilized in the medical literature to summarize available evidence on a research question. Like other studies, systematic reviews are at risk for bias from a number of sources. A systematic review should be based on a formal protocol developed and made publicly available before the conduct of the review; deviations from a protocol with selective presentation of data can result in reporting bias. Evidence selection bias occurs when a systematic review does not identify all available data on a topic. This can arise from publication bias, where data from statistically significant studies are more likely to be published than those that are not statistically significant. Systematic reviews are also susceptible to bias that arises in any of the included primary studies, each of which needs to be critically appraised. Finally, competing interests can lead to bias in favor of a particular intervention. Awareness of these sources of bias is important for authors and consumers of the scientific literature as they conduct and read systematic reviews and incorporate their findings into clinical practice and policy making.

Journal of Investigative Dermatology (2016) **136,** e109–e114; doi:10.1016/j.jid.2016.08.021

CME Activity Dates: October 22, 2016
Expiration Date: October 21, 2017
Estimated Time to Complete: 1 hour

Planning Committee/Speaker Disclosure: All authors, planning committee members, CME committee members and staff involved with this activity as content validation reviewers have no financial relationship(s) with commercial interests to disclose relative to the content of this CME activity.

Commercial Support Acknowledgment: This CME activity is supported by an educational grant from Lilly USA, LLC.

Description: This article, designed for dermatologists, residents, fellows, and related healthcare providers, seeks to reduce the growing divide between dermatology clinical practice and the basic science/current research methodologies on which many diagnostic and therapeutic advances are built.

Objectives: At the conclusion of this activity, learners should be better able to:
- Recognize the newest techniques in biomedical research.
- Describe how these techniques can be utilized and their limitations.
- Describe the potential impact of these techniques.

CME Accreditation and Credit Designation: This activity has been planned and implemented in accordance with the accreditation requirements and policies of the Accreditation Council for Continuing Medical Education through the joint providership of William Beaumont Hospital and the Society for Investigative Dermatology. William Beaumont Hospital is accredited by the ACCME to provide continuing medical education for physicians.
William Beaumont Hospital designates this enduring material for a maximum of 1.0 *AMA PRA Category 1 Credit(s)*™. Physicians should claim only the credit commensurate with the extent of their participation in the activity.

Method of Physician Participation in Learning Process: The content can be read from the Journal of Investigative Dermatology website: http://www.jidonline.org/current. Tests for CME credits may only be submitted online at https://beaumont.cloud-cme.com/RTMS-Nov16 – click 'CME on Demand' and locate the article to complete the test. Fax or other copies will not be accepted. To receive credits, learners must review the CME accreditation information; view the entire article, complete the post-test with a minimum performance level of 60%; and complete the online evaluation form in order to claim CME credit. The CME credit code for this activity is: 21310. For questions about CME credit email cme@beaumont.edu.

INTRODUCTION

Systematic reviews are comprehensive overviews of the existing evidence on a specific research question. If appropriate, they can include a pooled statistical summary of available data called a meta-analysis. Systematic reviews and meta-analyses are becoming increasingly prevalent in medical journals; a PubMed search using "systematic reviews" as a publication type filter in the *Journal of Investigative Dermatology*, *Journal of the American Academy of Dermatology*, *JAMA Dermatology*, and *British Journal of Dermatology* returned 7 results published in 2010 compared with 27 in 2015, although these figures may capture some narrative reviews as well. The results of systematic reviews can help guide clinicians, patients, and policy makers by providing more precise and comprehensive information than individual studies alone. They can also be used to identify gaps in knowledge and suggest areas for future research. A previous paper in the *Research Techniques Made Simple* series

[1]*Department of Dermatology, Warren Alpert Medical School, Brown University, Providence, Rhode Island, USA;* [2]*Division of Dermatology, University of Toronto, Toronto, Ontario, Canada; and* [3]*Division of Dermatology, University of Toronto and Women's College Research Institute, Toronto, Ontario, Canada*

Correspondence: Aaron M. Drucker, Department of Dermatology, Brown University, Box G-D, Providence, Rhode Island 02903, USA.E-mail: aaron_drucker@brown.edu

Abbreviation: PRISMA-P, Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol

SUMMARY POINTS

- It is important for authors of systematic reviews to:
  - Register a protocol before conducting the review and explain any deviations from it
  - Utilize the PRISMA-P and PRISMA guidance
  - Search comprehensively beyond the published literature
  - Assess risk of bias in included primary studies
  - Disclose competing interests.
- It is important for consumers of systematic reviews to be aware of those same issues when reading review reports and when interpreting the implications of their findings on clinical practice and policy.

discussed the methodology and utility of systematic reviews and meta-analyses in dermatology (Abuabara et al., 2012). In this article, we discuss the various types of bias that can occur in systematic reviews so that they can be avoided or acknowledged by review authors, and critically assessed by users of the dermatology literature.

REPORTING BIAS AND THE IMPORTANCE OF PROTOCOLS

Reporting bias refers to the selective dissemination of research findings based on the nature of the results (Kirkham et al., 2010). For example, the choice of review outcomes or included studies might be changed to highlight significant findings. The selective inclusion of outcomes or studies with more significant results after exploring the data will bias the results of the review toward positive findings.

To help identify and deter reporting bias, it is critical for systematic reviews to be conducted in accordance with a protocol written before beginning the review. As with other types of research, the protocol defines the research question—including the population, intervention or exposure, and outcomes of interest—and describes the methodology in sufficient detail to allow replication by others. To avoid a data-driven hypothesis, the research question should be formulated in advance based primarily on clinical relevance rather than knowledge of available evidence. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol (PRISMA-P) statement is a valuable evidence-based resource that defines the key content of a review protocol, including a description of search strategies and data sources, eligibility criteria, method of study screening and selection, primary and secondary outcomes, data extraction, and any planned analyses (Shamseer et al., 2015). Similarly, it is recommended that authors adhere to the comprehensive PRISMA guidance when actually preparing reports of systematic reviews (Moher et al., 2009). The PRISMA statement contains an evidence-based checklist of items to address in the manuscript itself and has been endorsed by many medical journals.

Public availability of the review protocol facilitates critical appraisal of the methods and identification of protocol deviations and selective reporting of results. It is important that protocols be prospectively registered online at PROSPERO—an online database of systematic reviews (http://www.crd.york.ac.uk/prospero/). Alternatively, protocols may be published in their entirety (as with Cochrane reviews). Subsequent publications of systematic reviews should state where the protocol was registered and where a copy of the protocol can be found. Protocol deviations do not necessarily lead to bias but must be explained in the Methods section of the systematic review report. For example, the search strategy might be modified if the results obtained from the original search were too broad or narrow. A recently published meta-analysis by Atzmony et al. (2015) concerning adjuvant

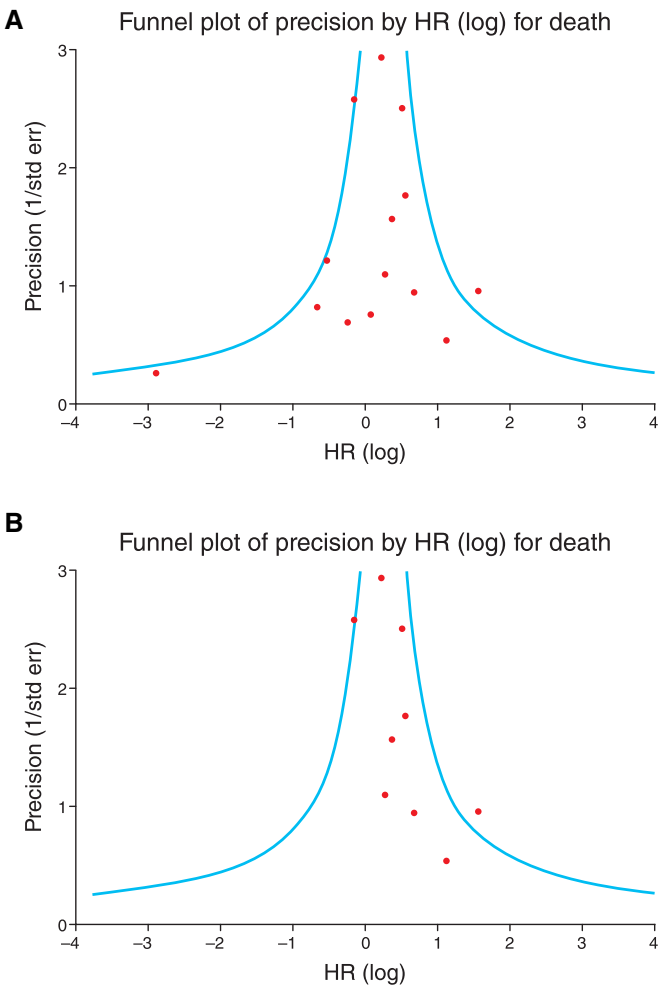

Figure 1. Representative examples of funnel plots. Funnel plots are scatter plots representing effect estimates on the *x*-axis compared with study precision (often the standard error of effect estimates) on the *y*-axis. (**a**) A symmetrical funnel plot adapted from a meta-analysis on the use of sirolimus in renal transplant recipients (Knoll et al., 2014). In this plot, the *x*-axis (log hazard ratio [HR]) is a proxy for effect estimates and the *y*-axis (standard error) is inversely related to the study sample size. The data points (red circles) each refer to a specific study. In a symmetrical funnel plot, the data points should be scattered symmetrically within the funnel (blue lines), suggesting a low risk of publication bias. (**b**) In this fictional plot (modified from Knoll et al., 2014), there is clear asymmetry within the funnel, with missing data points from unpublished trials in the lower-left portion of the funnel, suggesting a high risk of publication bias. Reproduced from Knoll et al., 2014 with permission from BMJ Publishing Group Ltd.

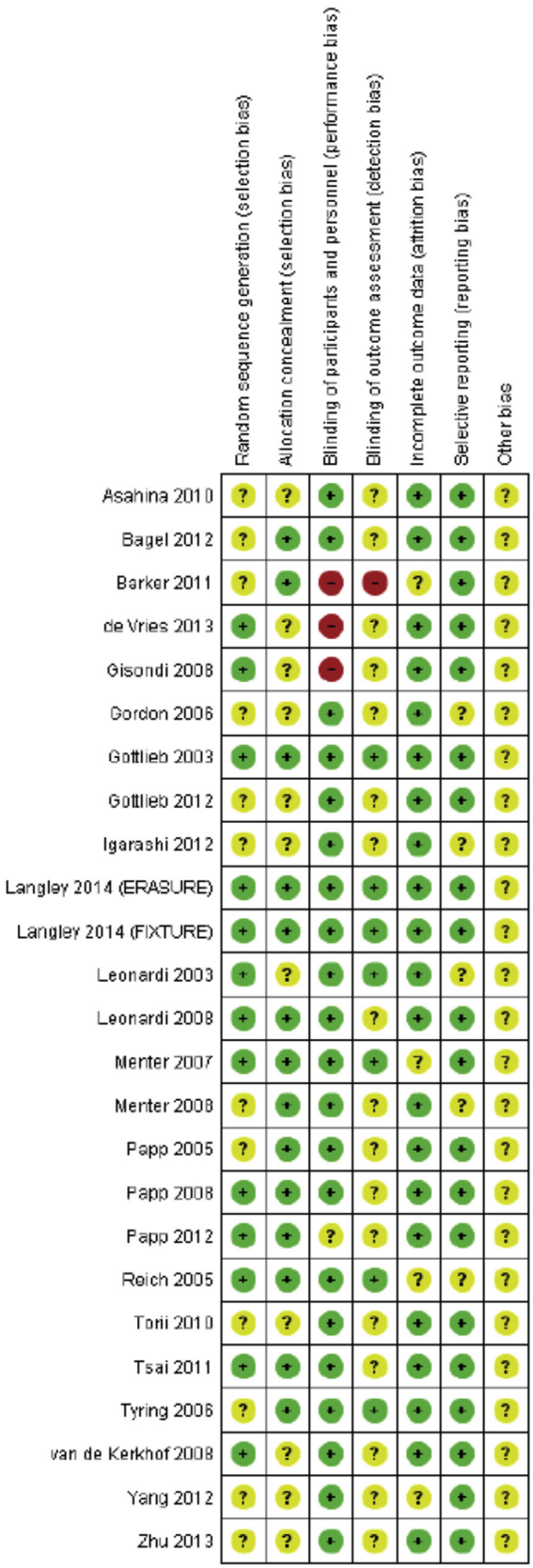

Figure 2. Tabular representation of risk of bias in individual studies. The authors of this systematic review on the efficacy of systemic treatments for psoriasis used the Cochrane risk of bias tool to assess potential sources of bias in included clinical trials, rating each as low (−), high (+), or unclear (?) risk of bias. Reprinted with permission from Elsevier from Nast et al. (2015).

therapy for pemphigus was reported based on PRISMA and was registered on PROSPERO (http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014014160).

EVIDENCE SELECTION BIAS

A key goal of a systematic review is to identify all relevant data to answer the research question. Data sources can include journal databases (e.g., PubMed), trial registries (e.g., clinicaltrials.gov), and direct communication with authors (Chan, 2012). Although electronic databases such as PubMed and EMBASE have made it much easier to identify published articles, systematic reviews are still prone to evidence selection bias from missed studies.

One source of this selection bias is publication bias. Substantial research has shown that only half of studies conducted are ever published. Statistically significant findings are more likely to be published and are published a year earlier, on average, than studies with nonsignificant findings (Dwan et al., 2013; Hopewell et al., 2007). Excluding unpublished, statistically nonsignificant data will bias a systematic review toward positive findings.

Although evidence selection bias due to nonpublication can be difficult to identify, there are a number of strategies that systematic reviewers can employ to search for all existing data (Chan, 2012; Liberati et al., 2009). Clinical trial registries, regulatory agency websites, and conference abstracts can be searched to identify unpublished studies or any outcomes that may have been selectively omitted from a study publication. The World Health Organization's clinical trials search portal (http://apps.who.int/trialsearch) is a useful tool to search across multiple registries. In a systematic review on the epidemiology of angiolymphoid hyperplasia, Adler et al. (2016) completed a comprehensive search that included traditional databases (PubMed and EMBASE) in addition to conference abstracts, Google Scholar, and manual searches of reference lists. Their search strategy was published as an appendix to the main report.

There are graphical and statistical methods that can be used to assess publication bias, though these have limitations. They rely on the assessment of the relationship between effect estimates (the magnitide of the exposure's effect on the outcome, such as the relative risk of infection between two interventions) and some measure of sample size for studies included in the systematic review (Higgins and Green, 2011). As a general rule, the precision of an effect estimate increases (and its confidence interval decreases) as the sample size increases. As a result, increased between-study variability in effect estimates is expected among smaller studies. Graphically, this can be represented by a funnel plot. Figure 1a shows a funnel plot modified from a meta-analysis on the use of sirolimus in renal transplant recipients (Knoll et al., 2014). In this plot, data points (representing individual studies) tend to scatter more horizontally in a symmetric funnel shape as the inverse of standard error (which is related to sample size) on the *y*-axis increases. This visual symmetry and funnel shape suggest a low risk of publication bias. This is in contrast to a fictional asymmetrical funnel plot shown in figure 1b, which suggests a higher risk of publication bias. Statistical tests of funnel plot asymmetry, such as Egger's test for

continuous outcomes, can help assess whether the association between effect estimates and standard error is statistically significant (Higgins and Green, 2011; Sedgwick and Marston, 2015). Although funnel plots and statistical tests are useful tools, they have limitations. Statistical tests have low power to detect asymmetry if there are less than 10 studies. Further, there are other causes of asymmetric funnel plots aside from publication bias, such as the inclusion of studies with heterogeneous patient populations, different study designs, or poor methodological quality (Sedgwick and Marston, 2015).

RISK OF BIAS IN PRIMARY STUDIES

Given that systematic reviews rely on data from other studies, the evidence in a systematic review is only as good as, or as free from bias as, the primary data sources. As such, each individual study included in a systematic review should be assessed for key sources of bias. Selection bias refers to the existence of systematic differences in baseline characteristics between the groups compared in a study. In randomized trials, selection bias can arise from inadequate generation of a random allocation sequence or inadequate concealment of allocations before group assignment. Detection bias arises from differences in outcome assessment due to knowledge of treatment allocation by unblinded outcome assessors. Performance bias refers to a systematic difference between groups in terms of how they are treated, or differences in the behavior of participants due to knowledge of the allocated interventions. Attrition bias refers to systematic differences in dropouts between groups. Finally, outcome reporting bias occurs when published trials selectively report only a subset of measured outcomes (Chan et al., 2014). A more detailed discussion of bias in primary studies can be found in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins and Green, 2011).

It is important that review authors report the methods used to assess the risk of bias in individual studies, as well as the findings of the assessment. Figure 2 presents an assessment using the Cochrane Collaboration's tool for assessing risk of bias in a systematic review of systemic treatments for psoriasis (Higgins et al., 2011; Nast et al., 2015). In this example, each study is graded as low (−), high (+), or unclear (?) risk of bias across different types of bias. The Cochrane risk of bias tool provides a domain-based qualitative description of critical areas of potential bias in clinical trials (Higgins et al., 2011). For meta-analyses, authors can conduct sensitivity analyses that exclude trials at high risk of bias to determine the effect on the results. Use of the Cochrane risk of bias tool is strongly recommended over the use of quality scales because the latter do not provide reliable measures of bias, and the summary scores they produce are difficult to interpret due to uncertainty over how each scale item should be weighted (Higgins and Green, 2011).

COMPETING INTERESTS

Systematic reviews conducted with ties to industry, particularly those funded by an industry sponsor, have the potential for bias in favor of the sponsor's product. Using Oxman and Guyatt's quality index for bias, Jørgensen et al. (2006) found that industry-sponsored reviews were generally of poorer quality than those conducted independently or sponsored by not-for-profit organizations. In addition, despite similar effect estimates, none of the industry sponsored reviews versus all of the non-industry-sponsored reviews expressed reservations in recommending use of the studied interventions (Jørgensen et al., 2006). To acknowledge the potential for bias related to industry involvement, all conflicts of interest must be disclosed by the authors of a systematic review, including who sponsored the study and what role the sponsor had in its design, conduct, and reporting.

Clinicians, patients, and policy makers sometimes rely on abstracts or summaries rather than the full systematic review report. However, misleading conclusions or "spin" often appears in summaries and abstracts of industry-sponsored systematic reviews (Yavchitz et al., 2016). Spin can also be used by academic authors who may overstate conclusions to increase the likelihood of their study being accepted for publication. Although spin may at times be obvious, it is not always readily apparent.

Spin in systematic reviews has been classified into three major categories:

1. Misleading reporting (e.g., not fully reporting the methods used to collect data)
2. Misleading interpretation (e.g., discussing nonsignificant results as if they were significant)
3. Inappropriate extrapolation (e.g., application of the study results to a patient population not actually studied in the systematic review) (Yavchitz et al., 2016).

It is important for authors to avoid spin when writing reports of systematic reviews and for readers to recognize spin by considering the full reports of systematic reviews in context before interpreting the conclusions.

OTHER QUALITY INDICATORS

Apart from sources of bias, there are a number of other quality-related issues to consider. When interpreting systematic review findings, authors should discuss the precision or uncertainty of a meta-analysis in terms of the 95% confidence interval of the summary effect estimate (Guyatt et al. 2011). It is also important for authors to evaluate the degree to which the results of primary studies included in the systematic review are consistent with each other. Heterogeneity refers to differences in results between primary studies that are greater than expected by chance alone (Higgins and Green, 2011). It arises from differences in various aspects of study design and conduct, such as the patient populations, interventions, outcome measurement methods, and quality. Heterogeneity can be quantified using the I^2 statistic (Higgins et al., 2003). A substantial degree of heterogeneity can be explored and addressed in a variety of ways. For example, the a priori selection of either a fixed- or random-effects model is an important consideration when conducting a meta-analysis with the potential for heterogeneity (Higgins and Green, 2011; Riley et al., 2011). A full discussion on the methodology and utility of meta-analyses in dermatology is available in a prior *Research Techniques Made Simple* article (Abuabara et al., 2012).

SUMMARY

Although systematic reviews and meta-analyses are invaluable for synthesizing available evidence, they are susceptible to

MULTIPLE CHOICE QUESTIONS

1. The protocol for a systematic review should…
   A. Be written and made publicly available before conducting the review
   B. Contain information on the sources of data that will be used
   C. Contain information on the outcomes that will be assessed
   D. Contain information on the criteria used to include and exclude studies
   E. All of the above

2. Publication bias occurs because…
   A. The peer review process takes too long
   B. Studies with statistically nonsignificant findings are less likely to be published
   C. Journals prefer to publish studies with nonsignificant findings rather than those with statistically significant findings
   D. Systematic reviewers change their outcome of interest after designing their protocol

3. Because they involve searches of the existing literature and pooling of multiple primary studies, systematic reviews…
   A. Are not prone to bias because they have a larger sample size than primary studies
   B. Are always able to find all available data on a topic
   C. Have their own sources of bias in addition to biases that exist in any primary studies
   D. Are easy to conduct and can be accomplished without significant effort

4. Which of the following is not a type of "spin"?
   A. Discussing limitations (e.g., explaining potential sources of missing data)
   B. Misleading reporting (e.g., not fully reporting the methods used to collect data)
   C. Misleading interpretation (e.g., discussing nonsignificant results as if they were significant)
   D. Inappropriate extrapolation (e.g., application of the study results to a patient population not actually studied in the systematic review)

5. Which of the following is true regarding competing interests?
   A. Authors of systematic reviews should disclose all potential competing interests
   B. If a nonindustry organization funded the conduct of a systematic review, but did not design, conduct, or write the review, then the funding does not need to be declared
   C. Although declaring competing interests is still important, funding from industry has been shown to have no effect on the results of systematic reviews
   D. Systematic reviews without industry funding never use "spin" when writing systematic review reports

multiple forms of bias. Authors of systematic reviews can minimize the risk of bias and promote transparency by registering and publishing the protocol before starting the review and by adhering to the PRISMA-P and PRISMA statements. It is important to explain protocol deviations and to search comprehensively for published and unpublished studies. Further, assessing the risk of bias in the included primary studies provides an indication of the quality of available evidence. Finally, review authors ought to acknowledge other sources of bias, including industry sponsorship. It is important for consumers of systematic reviews, including clinicians, patients, and policy makers, to be aware of these potential biases when reading systematic reviews and assessing the evidence they provide to address a clinical or policy question.

CONFLICT OF INTEREST

AMD is an investigator for Regeneron and Sanofi (no compensation received) and has received honoraria (speaker) from Astellas Canada.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to this paper. Teaching slides are available as supplementary material.

REFERENCES

Abuabara K, Freeman EE, Dellavalle R. The role of systematic reviews and meta-analysis in dermatology. J Invest Dermatol 2012;132:e2.

Adler BL, Krausz AE, Minuti A, Silverberg JI, Lev-Tov H. Epidemiology and treatment of angiolymphoid hyperplasia with eosinophilia (ALHE): a systematic review. J Am Acad Dermatol 2016;74:506–1.

Atzmony L, Hodak E, Leshem YA, Rosenbaum O, Gdalevich M, Anhalt GJ, et al. The role of adjuvant therapy in pemphigus: a systematic review and meta-analysis. J Am Acad Dermatol 2015;73:264–71.

Chan AW. Out of sight but not out of mind: how to search for unpublished clinical trial evidence. BMJ 2012;344:d8013.

Chan AW, Song F, Vickers A, Jefferson T, Dickersin K, Gøtzsche PC, et al. Increasing value and reducing waste: addressing inaccessible research. Lancet 2014;383:257–66.

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:e66844.

Guyatt G, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence-imprecision. J Clin Epidemiol 2011;64:1283–93.

Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.

Higgins J, Green S. Cochrane handbook for systematic reviews of interventions, version 5.1.0. The Cochrane Collaboration; 2011.

Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–60.

Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.

Hopewell S, Clarke M, Stewart L, Tierney J. Time to publication for results of clinical trials. Cochrane Database Syst Rev 2007;(2):MR000011.

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

Kirkham JJ, Altman DG, Williamson PR. Bias due to changes in specified outcomes during the systematic review process. PloS One 2010;5:e9810.

Knoll GA, Kokolo MB, Mallick R, Beck A, Buenaventura CD, Ducharme R, et al. Effect of sirolimus on malignancy and survival after kidney transplantation: systematic review and meta-analysis of individual patient data. BMJ 2014;349:g6679.

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. PLoS Med 2009;6:e1000100.

Moher D, Liberati A, Tetzlaff J, Altman DG. The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. BMJ 2009;339:b2535.

Nast A, Jacobs A, Rosumeck S, Werner RN. Efficacy and safety of systemic long-term treatments for moderate-to-severe psoriasis: a systematic review and meta-analysis. J Invest Dermatol 2015;135:2641–8.

Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. BMJ 2011;342:d549.

Sedgwick P, Marston L. How to read a funnel plot in a meta-analysis. BMJ 2015;351:h4718.

Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015;349:g7647.

Yavchitz A, Ravaud P, Altman DG, Moher D, Hrobjartsson A, Lasserson T, et al. A new classification of spin in systematic reviews and meta-analyses was developed and ranked according to the severity. J Clin Epidemiol 2016;75:56–65.