Methods Guidance

Number xx

Finding Evidence and Assessing for Reporting Biases when Comparing Medical Interventions: AHRQ and the Effective Health Care Program

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. xxx-xx-xxxx

Prepared by:
<Name> Evidence-based Practice Center
<City, State>

Investigators:

AHRQ Publication No. xx-EHCxxx
<Month Year>

This report is based on research conducted by the XXXXX Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. XXX-20XX-XXXXX). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this report.

Suggested citation: Finding Evidence and Assessing for Reporting Biases when Comparing Medical Interventions: AHRQ and the Effective Health Care Program. Methods Guidance No. NNN. (Prepared by the University of Ottawa and the Oregon Health and Science University Evidence-based Practice Centers under Contract No. XXXXXXX.) AHRQ Publication No.XXXXX. Rockville, MD: Agency for Healthcare Research and Quality. August 2012. www.effectivehealthcare.ahrq.gov/reports/final.cf

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies and strategies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

To improve the scientific rigor of these evidence reports, AHRQ supports empiric research by the EPCs to help understand or improve complex methodologic issues in systematic reviews. These methods research projects are intended to contribute to the research base in and be used to improve the science of systematic reviews. They are not intended to be guidance to the EPC program, although may be considered by EPCs along with other scientific research when determining EPC program methods guidance.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality. The reports undergo peer review prior to their release as a final report.

We welcome comments on this Methods Research Project. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by e-mail to epc@ahrq.hhs.gov.

Carolyn M. Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Director
Evidence-based Practice Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Task Order Officer
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Acknowledgments

The authors gratefully acknowledge the following individuals for their contributions to this project: <Acknowledgments>.

Key Informants

<Name>
<Place>
<City>, <ST>

Technical Expert Panel

<Name>
<Place>
<City>, <ST>

<Name>
<Place>
<City>, <ST>

Peer Reviewers

<Name>
<Place>
<City>, <ST>

<Name>
<Place>
<City>, <ST>

Contents

Tables

Figures

Boxes

Appendixes

Introduction

"Search for the truth is the noblest occupation of man; its publication is a duty" [Baronne Anne Louise Germaine de Staël-Holstein (1766-1817)].[1]

Systematic reviews attempt to identify, appraise and synthesize the available empirical evidence in order to minimize bias when representing the results of medical interventions and therapies. However, there is a growing recognition that often evidence is difficult to find because of decisions that are made about where, how, and when to publish the results of studies based on the findings of those studies. Notwithstanding, when unpublished data are actually available (for example as a result of legal action), reporting bias associated with suppression of unfavorable results has been fairly easy to detect.[2-5] A review by Song, et al. notes that the results of half of all clinical trials are never published. Other findings were that studies with positive or statistically significant effects tend to report greater treatment effect, tend to be published sooner and in higher impact journals than those with negative or nonsignificant effects, and that exclusion of non-English language literature may bias our understanding of treatment effects, particularly in the area of complementary and alternative medicine.[6]

Overview of Guidance

Since evidence syntheses depend on the published literature accurately representing what's known about medical therapies, reporting biases threaten the veracity of what we know. This document provides guidance on steps that authors of systematic reviews can take to reduce the error in the assessment of the effect of an intervention that arises from biases in the way that studies are published and reported.

The series of steps involved in searching for and identifying eligible studies for the review is lengthy and resource intensive. It involves searches that often turn up no additional studies, despite the searchers investment in time that can run into the hundreds of hours. Review teams are thus naturally reluctant to take on more searching than absolutely necessary. That said, in recent years it has become clear that the likelihood of finding a critical unpublished study or study data that changes key summary outcomes may be greater than we thought. For this reason, we are recommending searching these other sources for studies that might otherwise not have been identified. We understand that the number of potential sources for searching is large, and that the task of searching for unreported studies and data can never be considered "complete," because the "truth" is unknown.

Accordingly, we temper our recommendation for searching other sources with a recommendation to be selective and to choose the sources to be searched where it makes most sense. If a review concerns a drug used off-label, U.S. Food and Drug Administration (FDA) records will not contain effectiveness data for that indication, although they might well contain adverse effect data which could be useful across indications. In another example, if a condition is well-studied in another country (e.g., stroke trials in Japan), it may be a good idea to pay attention to the literature from that country and in that language. In a third example, given the fact that only 60 percent of randomized controlled trials (RCTs) described in conference abstracts reach full publication, and full publication is associated with results favoring the test intervention, then conference abstracts from the meeting most likely to publish trial abstracts is probably worth searching. That said, before searching, the systematic reviewers should check Cochrane's Master List of Journals being searched, and with the relevant review group, to make

sure this task hasn't already been done, with the findings available in the Cochrane Collaboration's Central Register of Controlled Trials.

The earlier guidance chapter by Relevo and Balshem[7] (referred to subsequently as Finding Evidence) provides guidance on the standard search for evidence. Here, we expand on that guidance and describe supplementary searches that should be considered as approaches to mitigating the effects of reporting bias. We describe the major data sources that should be considered when searching for unpublished studies and for published studies that are not likely to be identified through a search of the sources described in Finding Evidence. We discuss when those sources are likely to provide useful evidence and provide guidance on when searches of these sources should be considered.

We do not address the issue of multiple publication bias in this guidance. Multiple publication bias occurs when studies with significant or positive results are reported in multiple publications without citing the other reports of the same study. Instead we focus on providing guidance on identifying studies through the use of special searches, such as contacting authors, use of data from regulatory sites, use of protocols, hand searching, and the inclusion of non-English language literature, to reduce the likelihood of bias in estimates of effects of interventions.

Methods

Workgroup Composition

The workgroup for this chapter included 15 investigators and research associates from seven Evidence-based Practice Centers (EPCs) and the Agency for Healthcare Research and Quality (AHRQ). Nearly all workgroup members were authors of multiple systematic reviews with experience in addressing issues of reporting bias, and several have written extensively on the topic. A research librarian with several years experience in conducting searches for systematic and comparative effectiveness reviews was also a member of the workgroup. The topic was co-led by the Oregon and Ottawa EPCs. Project leadership involved establishing timelines, coordinating and scheduling conference calls, participation in subgroups, contributing to the writing of multiple sections of the guidance, and editing the overall guidance.

Guidance Development

We split the workgroup into two subgroups. A subgroup on Comprehensive and Special Searches focused on issues of finding all relevant published and unpublished literature. The second workgroup on Selective Outcome and Selective Analysis Reporting (SOR/SAR) focused on how to identify and assess the likelihood of biases arising from SOR/SAR. Each workgroup member participated in one or more subgroups.

The research librarian conducted a search for literature on topics related to reporting biases and compiled an EndNote library of relevant sources. Additional searches for literature were conducted at the request of the workgroups. The search identified more than 450 references spanning the period from 1959 through 2012.

The resulting guidance is based on empiric evidence, where available, and on experience and consensus where evidence was ambiguous or unavailable. Drafts of each subsection were first reviewed by the subgroup responsible for those sections. Subsequently a combined draft of both subsections was reviewed by all workgroup members and revisions made based on that review. The revised draft was then submitted for review by all EPC directors and others at the EPCs interested in providing comments, as well as by an Associate Editor of the Effective Health Care Program and the project Task Order Officer from AHRQ. We revised the guidance to address the major concerns of the EPCs and submitted a revised draft for external peer review and public comment. Comments from reviewers and potential edits were discussed by the workgroup both through conference calls and email. The document will be revised again based on peer review and public comment.

Background

Definitions and History

The Institute of Medicine has recently described reporting bias as "the greatest obstacle to obtaining a complete collection of relevant information on the effectiveness of healthcare interventions."[8] Reporting bias occurs when the dissemination and reporting of research results is influenced by the nature and direction of the findings. The selective publication of results—often those that are statistically significant ("positive") over nonsignificant ("negative") or null results—has been recognized for centuries.[9] Despite this, research was not undertaken to describe the size of the problem until about 50 years ago, when Sterling raised concerns that research yielding nonsignificant results was generally not published.[10] He confirmed his findings 35 years later in a second survey,[11] and to this day new articles continue to demonstrate the existence of sizable publication bias.[12-18] Box 1 describes several types of reporting biases that have been identified in the literature.

Box 1. Definitions of some types of reporting biases[19]

Publication bias
The publication or nonpublication of research findings, depending on the nature and direction of the results.

Time lag bias
The rapid or delayed publication of research findings, depending on the nature and direction of the results.

Multiple publication bias
The multiple or singular publication of research findings, depending on the nature and direction of the results.

Location bias
The publication of research findings in journals with different ease of access or levels of indexing in standard databases, depending on the nature and direction of results.

Citation bias
The citation or noncitation of research findings, depending on the nature and direction of the results.

Language bias
The publication of research findings in a particular language, depending on the nature and direction of the results.

Outcome reporting bias
The selective reporting, in published studies, of some outcomes but not others, depending on the nature and direction of the results.

Analysis reporting bias
The selective reporting, in published studies, of some analyses but not others, depending on the nature and direction of the results.

Reporting biases result both from the absences of complete studies from the body of literature and from the selective reporting of outcomes and analyses within individual study reports. While all publications necessarily select outcomes and analyses to report, outcome reporting bias (ORB) and analysis reporting bias (ARB) occur when outcomes are selectively reported (SOR) or data selectively analyzed (SAR)—typically in a *post-hoc* fashion—to favor a hypothesis.

An example of selective outcome reporting might be when a trial protocol indicates the primary outcome is evaluating an intervention's effect on increasing survival and the publication of the trial's primary results does not mention survivorship (for which there may have been no effect), but instead indicates that quality of life was the primary outcome, or reports results in a way that implies that quality of life was the primary outcome. Here the trial investigators have provided readers with information about certain outcomes and not others, and misrepresent outcomes as described in the protocol. Chan, et al. compared the contents of 102 trial protocols approved by the scientific ethics committees for Copenhagen and Frederiksberg, Denmark, during 1994 and 1995 with 122 subsequent publications.[20] They reported that in nearly two thirds of the trials there was a change in at least one primary outcome between the protocol and publication. The authors also reported that statistically significant outcomes had a higher likelihood of being reported compared with nonsignificant outcomes.

Selective analysis reporting operates in a similar manner. Here study authors may use selective cutoffs to dichotomize continuous outcomes or report selective time-point analyses when multiple time points were specified for analysis in the protocol.

SOR and SAR in published primary reports of individual studies may lead to biased interpretation of findings not only of individual studies but also of systematic reviews that include these studies.[21] Two studies provide empirical evidence of the effect of SOR and SAR on the pooled estimates of treatment effects.[22, 23] In addition, ORB and ARB may also operate at the systematic review level.[21-27]

Types of Selective Outcome Reporting and Selective Analysis Reporting

Selective outcome reporting and selective analysis reporting can be introduced at several points. At the protocol or conceptual stage of devising a study, bias occurs if investigators choose outcomes based on whether they will produce favorable results, rather than on their importance for clinical practice or policy decisionmaking. During results analysis, bias occurs if investigators decide to change their analysis (e.g., change in time point) in order to present favorable results. Additionally, results might be selectively reported (or withheld from reporting) to support competing interests. It may not be possible to determine whether some or all of these occur within a given study; this will depend on the extent of information available from other sources, such as the study protocol. Table 1 lists those types of SOR and SAR that could be identified and determined when assessing studies. Some of these constructs are also listed elsewhere.[28, 29]

Table 1. Types of selective outcome reporting and selective analysis reporting

Selective Outcome Reporting	Selective Analysis Reporting
• Partial reporting of outcomes (in other words, information is not sufficient to add the study to a meta-analysis) for example: Absolute or relative measure without either a confidence interval or a precise p value • Use of inexact p values (except p<0.01, which does not require more precision) • Narrative presentation of quantitative results (e.g., "significant" or "not significant") • Failure to report subgroups (prespecified or not) • Components of composite outcomes not reported • Change in the primary or secondary outcome • Omission of an outcome that was prespecified or for which clinical judgment suggests should have been prespecified • Addition of a new outcome that was not prespecified (excluding harms outcomes) • Use of different measurement techniques or scales • Change in the definition of categorical outcome • Incomplete specification of an outcome in the methods section of the publication or in other available sources	• Change in analytic method • Change in data types, for example, from dichotomous to continuous • Change in effect measure • Change in assumption of data distribution • Change in time points for analysis • Early stopping for benefit • Post hoc subgroup and sensitivity analyses • Final compared with change from baseline analyses • Selectively reporting the first period results in crossover trials

Sources of Evidence

Institute of Medicine (IOM) standard 3.2 requires those conducting systematic reviews to "take action to address potentially biased reporting of research results."[8] This section discusses the various sources of data discussed in the IOM report, how they can be used in the search for evidence, and provides empirical evidence of their value as sources of information both on unpublished studies and of unpublished data in published studies, as well as evidence that excluding evidence from these sources can lead to biased effect estimates.

Study Registries

Study registries are publicly available databases or platforms, commonly Web-based, in which research studies are catalogued. In the last 5 years, several trial registries have evolved into data repositories of key elements of the trial protocols, including outcomes and/or their summary results. Trial registries can serve as a resource both for identifying unpublished studies and for identifying unreported outcomes in published studies.

The U.S. Food and Drug Administration Modernization Act (FDAMA) of 1997[30] mandates the registration of clinical trials that evaluate the efficacy of drugs for serious or life-threatening diseases and conducted under an Investigational New Drug Application. Beginning in 2005, the International Committee of Medical Journal Editors (ICMJE) required prospective trial registration as a precondition for publication.[26] Ongoing trials have to be registered by September 2005.[31] The U.S. Food and Drug Administration Amendments Act (FDAAA) of 2007[32] further requires that trials already in progress be registered on ClinicalTrials.gov by December 2007 and that researchers post a summary of basic results within a year of completion of data collection or within 30 days after the FDA first approved the drug[33] (See Table 2 for a description of some of the registration and reporting requirements established by the Act). ClinicalTrials.gov, launched in 2000 to comply with FDAMA, currently contains over 120,000 trials sponsored by the National Institutes of Health, other Federal agencies, and private industry.

Studies listed in the database are conducted in all 50 States and in 179 countries.[34] Appendix A describes the data elements available from ClinicalTrials.gov.

The World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) was established in 2005 as a portal that imports trial registration data from clinical trial registries around the world including ClinicalTrials.gov. It contains more than 180,000 records for nearly 170,000 trials, including records for more than 60,000 trials conducted in the United States.[35] Appendix B describes the data elements available from ClinicalTrials.gov.

Non-randomized studies can occasionally be found in study registries. Several trial registries, including ClinicalTrials.gov, ISRCTN/ControlledClinicalTrials, ANZCTR (Australia/New Zealand), Clinical Trials Registry-India, UMIN Clinical Trials Registry (Japan), and the Chinese Clinical Trials Registry, allow registration of observational studies, with observational studies representing 17 percent of all studies registered in ClinicalTrials.gov in the year 2010.[36] However, the utility of these external sources of registry data for identifying or minimizing reporting bias associated with observational studies has not yet been evaluated. There is growing interest in registration of observational studies, especially prospective observational studies,[36-38] although some have suggested that requirements to register observational studies might actually impede, rather than advance scientific discovery because serendipity, exploration and chance findings will be lost.[39, 40]

Table 2. Registration and reporting requirements of the U.S. Food and Drug Administration Amendments Act, Section 801[a] (reprinted with permission from Wood 2009[41])

Type of Requirement	Type of Trial	Deadline for Reporting	Type of Data	Effective Date
Registration	Applicable clinical trials of drugs or biologics and devices regulated by the FDA[b]	No later than 21 days after enrollment of first participant	- Summary protocol; population, study design, outcome measures - Recruitment information - Location and contact information	Dec. 26, 2007
Basic results reporting	Applicable clinical trials of approved drugs and biologics and cleared or approved devices regulated by the FDA[b]	No later than 1 year after completion date; delayed submission is permitted in some cases[b]	- Demographic and baseline characteristics of participant sample - Participant flow - Primary and secondary outcomes - Certain agreements regarding dissemination of results information	Sept. 27, 2008
Adverse events reporting			- Serious events - Frequent events	Sept. 27, 2009
Expanded results reporting	Examples include applicable clinical trials of unapproved drugs or biologics regulated by the FDA[b]	Examples include extension of submission date, up to 18 months after completion date, and reconsideration of timing and requirements for submitting updates[c]	Examples include technical or lay summaries and complete protocol or other information necessary to evaluate results	Sept. 27, 2010

Abbreviations: FDA, U.S. Food and Drug Administration.

[a] Information on trial registration, basic results reporting, and adverse events e-reporting is available at http://prsinfo.clinicaltrials.gov/definitions.html and at http://prsinfo.clinicaltrials.gov/fdaaa.html. The requirements for expanded results have not yet been defined.

[b] According to the FDA Amendments Act, an "applicable clinical trial" is generally one that has at least one trial site in the United States. Section 801 excludes phase 1 drug trials and "early feasibility device trials." All applicable clinical trials of devices must be submitted, but only trials of devices previously cleared or approved are posted. Note that the ICMJE and the WHO require registration of all clinical trials for drugs and devices, regardless of phase.

[c] According to the FDA Amendments Act, "completion date" refers to "the date that the final subject was examined or received an intervention for the purposes of final collection of data for the primary outcome, whether the clinical trial concluded according to the prespecified protocol or was terminated."

Empirical Findings on the Value of Searching Study Registries

Despite registration requirements more than half of the trials that reported start dates with their registration were registered late[33] and only 12 to 22 percent of trials posted results within one year of completion.[33, 42] The number of unregistered trials and those with missing results is unknown, as is the accuracy of the data submitted.[25] Compliance with the FDAAA mandatory reporting requirement of trial results is low: within one year of study completion, only 22 percent of 738 trials were compliant.[42] In a review of a sample of trials registered with the ICTRP between June 2008 and June 2009, Viergever and Ghersi[43] found that over half of the trials were registered after the date of first enrolment and that contact information was available for 94 percent of nonindustry funded and for 54 percent of industry funded trials. Compliance with the requirement to post results for both industry and nonindustry sponsored studies at ClinicalTrials.gov is also poor.[44] The proportion of registries with adequate reporting of trial methodology ranged from 1.4 percent (allocation concealment) to 66 percent (primary outcomes)

in a study of ClincialTrials.gov and six other registries supported by the WHO search portal ICTRP.[45]

In a study of National Institutes of Health funded trials registered in ClinicalTrials.gov, Ross, et al.[46] found that fewer than half the trials were published in a peer reviewed journal indexed in Medline within 30 months after trial completion. In an earlier study Ross, et al.[47] found that only 46 percent of all completed studies registered in ClinicalTrials.gov had been published, and that even when published, fewer than half of the registrations included a citation to the published report. Wieseler and associates compared journal publications, clinical study reports submitted to regulatory agencies, and trial registry information and noted that study information was most comprehensively reported in regulatory submissions with registry and publications complementing each other.[48]

Although study registration and the reporting of study results remains incomplete and may be delayed, trial registries can still help to identify both unpublished studies and unpublished outcomes in published studies.[21, 44, 47, 49-52] Dwan, et al.,[21] in their systematic review of the empirical evidence of study publication and outcome reporting bias, included studies of cohorts of trials examining discrepancies between trial registry entries and associated protocols and publications. Several discrepancies were noted – differences in reporting of sample size calculations (84 percent) and methods of allocation concealment (6 percent), handling of missing data (80 percent) blinding (67 percent), and primary outcome analysis (60 percent). Six other studies have shown similar discrepancies between trial registries and subsequent publications in reporting efficacy outcomes and adverse events (e.g., primary outcome omission, upgrading from secondary to primary outcome, new primary outcome introduction, underreporting of recurrent and low grade adverse events, incomplete description of adverse events, and tendency for reporting of statistically significant results favoring test drug).[17, 44, 47, 49, 50, 52]

Regulatory Documents

Reviews of Drugs compared with Devices

Drugs and devices are both regulated by the FDA. However, the regulatory requirements and the approval processes for drugs and devices can be quite different.[53] These differences, described below, limit the usefulness of searches of the FDA for information about effectiveness studies on medical devices.

Drug Approval Process

Manufacturers are required to submit a New Drug Application (NDA) to the FDA for all new drugs for which approval for marketing in the United States is sought. The FDA Center for Drug Evaluation and Research (CDER) reviews the clinical and preclinical data for the proposed indication and makes a determination of approval status. Findings of those reviews are included in a number of FDA documents.

While there are often dozens of documents and tens of thousands of pages produced during the course of the review, the two documents of most relevance to those conducting systematic reviews are the Medical Reviews (sometimes referred to as Clinical Reviews) and the Statistical Reviews. The Clinical Review is a comprehensive summary and analysis of the clinical data submitted in support of a marketing application and includes the FDA reviewer's assessment of and conclusions about: 1) the evidence of effectiveness and safety under the proposed conditions of use; 2) the adequacy of the directions for use; and 3) recommendations on regulatory action based on the clinical data submitted by an applicant. The Statistical Review describes key

statistical issues and findings that affect conclusions regarding the demonstration of efficacy/safety. It summarizes and discusses the reviewer's analyses, the extent of evidence in support of claims, and statistical issues that may affect the conclusion on efficacy and/or safety, and is based on a review of individual studies as well as on the collective evidence. In addition to the primary endpoint analysis, the statistical reviewer may also address secondary or subgroup analyses if these are deemed important.

Drugs@FDA, a web-based, searchable database of information about FDA-approved brand name and generic prescription and over-the-counter human drugs and biological therapeutic products, while challenging to use, provides access to Medical and Statistical Reviews (see Appendix C). (http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm)

Device Approval Process

Medical devices are regulated by the FDA Center for Devices and Radiological Health (CDRH), and while all devices must comply with regulations regarding good manufacturing practices, proper labeling, adequate packaging, and registration with the FDA, most devices are approved through a process that is much less demanding than that required for drugs and which, for most, does not require trials demonstrating safety and efficacy.[53] Prior to 1976 medical devices were not required to be registered with the FDA or to follow quality control standards prior to marketing, and have come to be known as predicate devices. Since 1976 devices are classified into one of three categories depending on their perceived level of risk. Class 1 devices are those considered to have the lowest level of risk and include devices such as such as tongue depressors and band-aids. Class II, which includes devices such as forceps and surgical lasers are considered to pose a greater level of risk. Class III devices are devices that support or sustain life, such as drug-eluding stents and pacemakers, and are considered to have the highest level of risk for injury or illness. Only Class III devices go through a process known as a Premarket Application (PMA) that is more similar to the process required for drugs, and requires a demonstration of sufficient scientific evidence to demonstrate safety and efficacy for the intended use. However, only about 2 percent of all devices are approved through the PMA process.

Empirical findings on the value of searching for regulatory documents

Relatively few studies have looked at the impact of including information from regulatory documents on the conclusions of comparative effectiveness reviews. Reviews of the use of FDA documents have found that inclusion of unpublished studies from FDA documents may reduce the estimate of effect found in published studies;[54] that FDA documents suggested an elevated risk of harms not acknowledged in FDA advisory committee recommendations;[55, 56]; that prompt analysis of data available to the FDA can identify harms not identified in the published literature;[57] that publication is associated with positive outcomes;[3] but that the highly selective nature of the populations included in the unpublished trials raise questions about the applicability of those findings to actual clinical practice.[58] Similarly, a review of published and unpublished data provided to the British Medicines and Healthcare products Regulatory Agency found that while published data indicated that benefits of the study drugs outweighed their risks, that the inclusion of unpublished data suggested that risks outweighed benefits for all but one of the drugs reviewed.[5]

Rising, et al.[17] compared publications with data submitted to regulatory agencies and found additional and omitted outcomes and reporting of different statistical analyses in the published versions. An updated Cochrane systematic review on oseltamivir for preventing and treating

influenza incorporated previously unpublished data obtained from regulators.[59] The authors found evidence of reporting bias in trial publications, and conclusions changed such that the drug could no longer be considered effective. Hart et al.[22] reanalyzed 42 meta-analysis of nine drugs with additional, unpublished data obtained from the FDA. Lower drug efficacy was found in 46 percent of reanalyses, identical efficacy in 7 percent, and greater efficacy in 46 percent. Harms were underestimated when the meta-analysis was restricted to published data.

However, even when available, FDA reviews can be difficult to find and use. O'Connor found that the search engine could fail to find a review even when using the application number, and noted that reviews are difficult to navigate, generally being quite long with inadequate or incorrect tables of contents.[60]

Study Protocols

A clinical study protocol is a document that provides details of the study plan and organization and is written prior to the start of subject recruitment and data collection. Protocols include information on study rationale, objectives, methodology (design and statistical approaches), types of participants (i.e., inclusion and exclusion criteria), treatments, clinical procedures, ethical considerations, and the duration of the study. (http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6_R1/Step4/E6_R1__Guideline.pdf, accessed on April 13, 2012; http://clinicaltrials.gov/ct2/info/glossary, accessed on April 13, 2012).[61]

Study protocols and related information can be located and accessed from several sources such as study authors, industry registries, trial registries, Websites of relevant agencies (e.g., ClinicalTrials.gov, canadatrials.com, controlled-trials.com, and World Health Organization International Clinical Trials Registry Platform), and through documents made public as a result of litigation. Also, several peer reviewed medical journals (e.g., *The Lancet*, Biomed Central, *Trials*) publish study protocols. *The Lancet* began publishing protocols of randomized trials in 1997 and extended this to observational studies in 2001.[62, 63] BioMed Central began publishing protocols for a variety of study designs in 2001.[64] In 2006, the journal *Trials* was launched and has accepted study protocols from the outset.[65]

Empirical Findings on the Value of Searching for Protocols

Several recently published empirical studies comparing protocols and published reports of individual trials for consistency and completeness of outcomes and analyses [20, 52, 66, 67] provide evidence of outcome reporting bias in published reports of individual RCTs. Dwan, et al. published two systematic reviews that summarize these findings.[13, 21] These studies report a high prevalence of unreported or incompletely reported outcomes. Outcomes with a statistically significant difference were more likely to be reported than outcomes associated with a nonsignificant difference (OR [odds ratio] 2.4, 95% CI [confidence interval], 1.4 to 4.0).[20] The primary outcomes specified in the protocols were either changed to secondary (and a new primary outcome was introduced), or omitted from the subsequent publication.[20, 52, 66, 67] In a review of study protocols examined as part of a litigation against Pfizer and Parke-Davis regarding off-label use of gabapentin, published primary outcomes differed from those described in the protocol in 8 of 12 reported trials and all changes between what was specified in the protocol and what was later published led to a more favorable presentation of the efficacy of gabapentin for unapproved indications.[68] However, finding protocols can be challenging. Hartling, et al. in their systematic review attempted to inform their study risk of bias assessments

by additionally retrieving protocols for 42 of 107 trials. No restrictions such as trial country of conduct or year of publication were employed. The yield was low (protocols could be obtained for just 12 percent of studies), with protocol retrieval adding 50 percent more time to risk of bias assessment.[69]

Conference Abstracts and Proceedings

Authors frequently present, in oral or poster form, interim or full study results at professional meetings. Often, meeting submissions are collated as a catalogue of abstracts.

Empirical Findings on the Value of Searching Conference Abstracts and Proceedings

In a review of findings initially presented as abstracts at European General Practice Research Network meetings from 1999-2002 and 2005-2006, Royen, et al. found overall 45 percent of the presentations to have been subsequently published, with abstracts from the 2005 to 2006 meetings having only a slightly higher publication rate (43 percent for the period 1999–2002 and 47 percent for the period 2005–2006).[70] Similarly, Scherer, et al. found that fewer than half of all abstracts were published in full, and that positive results were positively associated with full publication.[71] Tam and Hotte[72] compared a subset of phase III trials presented at the 2000 American Society of Clinical Oncology Annual Meeting with their subsequent full publication (by May 2006). Of 55 abstracts that were subsequently published, the primary endpoint was stated in 34 percent of abstracts compared with 100 percent of publications. Primary and secondary endpoints, primary endpoint results, statistical analysis, and statistical significance of the primary endpoint were frequently not clearly described in the abstract. For abstracts that were clearly described, primary endpoints were identical in 90 percent of cases; statistical significance of the primary endpoint and conclusions were identical in 89 percent and 91 percent of cases, respectively. The primary endpoint results differed by more than 5 percent in 42 percent of abstract-to-publication comparisons. However, abstracts and proceedings frequently report only preliminary results, which may not accurately represent what was found once all data were collected and analyzed.[73-75]

Grant Databases

Several grant databases allow for analysis of registration and publication status of all United States Federally funded studies (Appendix D).

The Federal Research Portfolio Online Reporting Tools (RePORT) database, the largest United States based grant database, provides several downloadable and analyzable data elements, including start and end dates, names and affiliations of principal investigators, financial information about the grants, and grant titles and project abstracts. The RePORT database does not include variables indicating study registration or participant recruitment status, rendering it difficult to determine if the study has been completed.

In addition, the current practice of posting all publications that mention a grant complicates attempts to determine a study's publication status. The RePORT website warns that articles posted on the site "are associated with projects, but cannot be identified with any particular year of the project or fiscal year of funding. Some publications will be inadvertently linked to the wrong grant or missing altogether." Most published articles include several grant numbers, and each grant project includes links to several articles. Published article titles and abstracts often differ from descriptions of the grants.

Empirical Findings on the Value of Searching Grants Databases

Empirical evidence shows low registration rates in clinical trial registries for federally funded trials.[76, 77] Recent studies that have examined the registration and publication of National Institutes of Health (NIH) funded studies have found poor availability of protocols and study results.[76, 78] The analysis of NIH funded pediatric trials demonstrated that only 33 percent were registered and only 53 percent were published.[76] The analysis of NIH funded therapeutic studies for female urinary incontinence found that only 6 percent were registered.[78] Published studies (94 percent of all NIH funded) mentioned the NIH grant numbers but did not necessarily report study results.[78]

We found no studies comparing the protocols of registered NIH funded studies with published results to evaluate deviations from the protocol and selective outcome reporting.

Contacting Authors to Identify Unpublished Studies

The completeness of reporting of individual studies (and systematic reviews themselves) is often suboptimal. Authors of a study may not have reported all of the outcomes specified in study protocols, may not have completely described the type of participants included in their study, or may have provided published analyses at the aggregate level when analyses were also done for subpopulations. Contacting study authors may be useful for obtaining missing or unreported outcomes, obtaining outcomes in a format suitable for meta-analysis, or to clarify potential errors or unclear results. Contacting authors might also provide additional information regarding study methods that may prove helpful in rating study quality.

Empirical Findings on the Value of Contacting Authors

There are few papers examining the utility of contacting authors in the context of conducting a systematic review. Mullan, et al. reviewed 147 published systematic reviews, of which 54 were Cochrane reviews. These reviews were published in high impact factor journals. The researchers reported that 46 (50 percent) of the traditionally published reviews and 46 (85 percent) of the Cochrane reviews reported contacting study authors.[79] Missing data was the most common reason for contacting study authors.

In a recent systematic review of the literature on methods for obtaining unpublished data, Taryn and colleagues found that, in general, requests to authors for clarification about study methods were more likely to be successful than requests for missing data about study results. . While contacting authors by email seems to result in the greatest response rate with the fewest number of attempts and the shortest time to respond, they also found that there is no consistent evidence about what approaches work best.[80]

Three studies not considered in the Taryn review assessed whether contacting authors for more information adds substantive information. Kyzas and colleagues[81] found that contacting authors (with second attempt at 2 months) and obtaining additional data (11 studies; 996 patients) changed results from statistically significant (RR [relative risk] 1.23, 95% CI, 1.03 to 1.47; 31 studies; 2,392 patients) to not significant (RR 1.16, 95% CI, 0.99 to 1.35, p=0.06; 3,388 patients). Taryn and colleagues noted, however, that response rates do not seem to be influenced by the number of requests.[80]

Chan, et al.[66] compared trial protocols with their published versions for 48 relatively large randomized studies funded by the Canadian Institutes of Health Research (1990−1998), the Canadian governmental funding agency. Eighty-eight percent of the 48 trials measuring efficacy and 62 percent of 26 trials measuring harms had at least one unreported outcome. They surveyed

authors, and of 43 respondents, 80 percent denied that any outcomes were unreported. When study authors were provided with a list of unreported outcomes at 6 weeks after the initial query, 37 respondents (77 percent) provided some details about the unreported outcomes. Kirkham,[23] in evaluating trials included in a cohort of Cochrane reviews for selective outcome reporting, contacted authors of 167 trials for additional information and received a response from only 39 percent of authors in 3 weeks. They were able to confirm and obtain reasons as to whether outcomes were measured and not analyzed or just not measured. The authors observed similar response rates for trials at high and low risk of suspected outcome reporting bias. It is not known how generalizable the above response rates are, particularly given that some reference older trials when authors were not as aware of such biases.

Handsearching

Handsearching refers to manually scanning print journals to identify relevant studies not retrieved by electronic bibliographic databases Not included within this definition of handsearching are reviews of reference lists and citation tracking, which are other methods for identifying potentially relevant citations. Handsearching may also be valuable for identifying studies published only as conference abstracts, since these are often published as journal supplements that are not included in electronic databases. Examples of situations in which relevant studies may be included in an electronic database but not well indexed include newer interventions that have not yet been assigned Medical Subject Headings (MeSH), and when systematic reviews address complex interventions or evaluate topics such as harms or subgroup effects that may not be indexed well.

Empirical findings on the Value of Handsearching

Less than a third of the world's medical journals are routinely indexed in the major electronic databases.[82] A Cochrane systematic review found that handsearching identified more relevant randomized trials (92 to 100 percent) than searches based on single electronic databases (range 49 to 77 percent).[83] However, more sensitive search strategies such as the Cochrane Highly Sensitive Search Strategy identified 80 percent of relevant randomized trials, or nearly as many as were found by handsearching. This systematic review did not compare the yield of handsearching with searches based on two or more electronic databases, or handsearching compared with searches on electronic databases, reference list reviews, and other supplemental methods, such as peer review suggestions. It also did not evaluate the yield of handsearching for nonrandomized intervention studies or studies of diagnosis or prognosis. One study found that handsearching for studies of diagnostic test accuracy of 18F-fluorodeoxyglucose positron emission tomography-computed tomography did not yield additional studies compared to database searching.[84]

Handsearching is time-consuming and resource intensive. Although no study has evaluated differences in estimates of effects when handsearches are conducted in addition to electronic database searches and other supplemental methods, the value of handsearching probably varies depending on the topic of the systematic review. The yield of handsearching is likely to be higher when relevant studies are published in journals that are not indexed in electronic databases, or in journals that are indexed in electronic databases but indexing is suboptimal, associated with a significant lag time, or published as a journal supplement.[85] Studies that may be less likely to be included in standard English-language electronic databases include older

studies, studies of complementary and alternative interventions, and non-English language studies.

Searching for Non-English Language Literature

Although most of the more significant medical literature is indexed in the major bibliographic databases such as MEDLINE and EMBASE, there is still a considerable amount of relevant and important literature published in non-English language journals that are not indexed by these databases. Identifying non-English language articles published in these journals may require a search of additional databases such as Global Index Medicus published under the auspices of the World Health Organization.

Empirical Findings on the Value of Searching the Non-English Language Literature

A Medline search of all publications from 2000 to February 3, 2011 found that of 6,574,939 citations, 90 percent were published in English. Table 3 shows the number and frequency of publications in other languages with at least 1 percent frequency.

Table 3. Percentage of publications from Medline in various languages (1996-2011)

Language	N	Percent
Total	**6,574,939**	**100%**
English	5,926,763	90%
Chinese	109,658	1.7%
French	97,752	1.5%
German	88,191	1.3%
Japanese	73,657	1.1%
Russian	71,583	1.1%
Spanish	71,281	1.1%

Based on a review of recent CER reports with final or draft documents downloadable from the AHRQ Web site, most (71 percent) EPC reports restricted literature searches to English language publications. Thus, EPC reports may be at risk of selection bias based on language, and may not be consistently following Institute of Medicine Standards for Systematic Reviews (Standard 3.2.6).

Empirical evidence, however, has not shown consistent findings regarding language bias. For example, investigators in Germany may be more likely to publish their negative results in German language publications and their positive results in English language publications,[86, 87] and almost all Chinese acupuncture trials published in Chinese report positive results.[88] Numerous other studies, however, have found that excluding non-English publications may not have an impact on the conclusions in systematic reviews.[89-95]

Information from Searches of the World Wide Web

Nearly all searches for evidence today, including searches for regulatory documents, registries, indexed literature, etc. are conducted on the web. In this section we take the phrase "search the world wide web" to mean using standard web search engines such as Google or Google Scholar, to supplement searches of specific web sites, such as the FDA web site

Drugs@FDA.com or ClinicalTrials.gov, or searches of proprietary databases such as MEDLINE and EMBASE.

Empirical Findings on the Value of Searching the World Wide Web

Several studies have compared the citation counts resulting from searches of Web of Science, Scopus, SciFinder, and Google Scholar.[96-99] All found considerable variation in the resulting citation counts. Kulkarni found that Google Scholar and Scopus retrieved more citations than Web of Science; that Scopus retrieved a greater proportion of non-English citations, but that Web of Science retrieved more citations from articles, editorials, and letters.[97] Li noted that Web of Science provides coverage back to 1990 while Scopus only provides complete coverage back to 1996, but found that Scopus provides better coverage of clinical medicine than Web of Science.[96]

Guidance on Assessing for Selective Reporting of Outcomes and Analyses

This section begins by providing guidance on when an expanded search, beyond the standard search described in the guidance on Finding Evidence,[7] may be appropriate. It then explains how the risk of ORB and ARB can be assessed and clarified once information on a study has been retrieved. The proposed assessments of ORB and ARB specifically reflect a study level *risk* (potential) for bias as it applies to the review, not the actual bias in the study (which may or may not be present). For example, authors may be genuinely limited by journal word count restrictions and hence report some outcomes in narrative form or omit them altogether. Such omissions would not necessarily result in biased effect estimates, unlike omissions related to the desirability of certain results. Because the intent of authors cannot be known by systematic reviewers, a thoughtful assessment of the risk of outcome and analysis reporting bias is required.

The review stage when grey literature is used for ORB and ARB assessments may vary across reviews. For example, when reviewers have searched trial registries, contacted authors, obtained relevant documents from industry, and acquired FDA documents up front as part of their standard review search strategy and used the search output to identify studies for which no published report was found (publication bias), they may have simultaneously identified unpublished study data and protocol details for published studies included in their review. As we recommend below, all information for a study should be examined together for risk of bias assessment and data extraction. In such a situation, the risk of ORB and ARB may be assessed without further searching or additional clarifications from unpublished sources of study information. Alternatively, when the primary search was restricted to published studies, reviewers might want to search and cross-check against those same sources while conducting ORB and ARB risk assessments.

Principles for Assessing Outcome Reporting Bias and Analysis Reporting Bias

Outcome Level Assessment

The risk of selective outcome and analysis reporting bias is an outcome-level assessment, as opposed to a study-level assessment. ORB and ARB may differ among outcomes because the decision to selectively present or omit outcomes or their analyses will depend directly on the results that were obtained for a given outcome. Similarly, risk of performance bias (e.g., blinding or masking of participants and providers) and detection bias (e.g., blinding of outcome assessors) entail outcome-level assessments, while selection bias (e.g., allocation concealment) is a study-level assessment.

Assess Important Outcomes Determined *a Priori*

Of outcomes of interest to the review, we suggest restricting ORB and ARB assessments to those outcomes that will be graded for their strength of evidence according to guidance provided by the EPC Program.[100] Gradable outcomes are those determined *a priori* during the topic refinement phase and reported in the protocol to be important for healthcare decisionmaking. We make this recommendation for practical reasons, given the volume of outcomes that can be included in an EPC systematic review. Review authors should evaluate reporting bias for their

prespecified gradable review outcomes irrespective of whether those outcomes were designated as primary or secondary in the study.

Assessment of Outcome Reporting Bias and Analysis Reporting Bias for Benefits and Harms

In general, reporting bias in trial publications takes the form in which benefits are over reported and harms under reported.[49, 101] ORB and ARB for harms can be addressed similarly to beneficial outcomes. However, in rare cases, it is possible that a serious harm was identified during the evidence synthesis process and was not previously identified for grading the strength of evidence; a *post hoc* decision may then be made to assess ORB specifically for that outcome.

Composite Outcomes

Reporting only of composite outcomes may be an indicator for the presence of ORB or ARB. A common example in cardiovascular research is the composite outcome of vascular death plus nonfatal myocardial infarction plus nonfatal stroke. Composite effects could mask the effects corresponding to individual components; we cannot assume the individual components have effects equal to the composite.[102] Studies that report composite outcomes should also provide results for the component outcomes.

Additional Considerations

Outcome and analysis reporting bias should be assessed comparing treatment effects on outcomes in all available reports of the same study (one or more articles, abstracts, results posted in clinicaltrials.gov, and FDA reviews) including their protocols (published protocols, protocols posted in clinicaltrials.gov, and methods sections in the articles). In general, systematic reviewers should recognize that studies that do not investigate or report outcomes of interest to the review may be susceptible to SOR or SAR, and so should not exclude such studies from the review.

Because of the potential impact on effect estimates, reporting bias should be cautiously assumed to exist even if authors cannot determine its direction and magnitude.

Identifying Selective Outcome Reporting and Selective Analysis Reporting in Included Studies

Above we described the various sources of information on study outcomes and analyses and the empirical evidence on the accuracy, completeness, and feasibility of using these sources to identify and characterize SOR and SAR. In this section, we provide guidance based on that evidence, and on expert opinion when the evidence base is insufficient. Our recommendations are likely to be revisited as new or more robust evidence emerges.

Appendix E compiles and summarizes the guidance discussed in the following sections.

The Initial Search for Evidence

The evaluation of the literature for selective outcome and analysis reporting begins with the search for evidence. The goal of the search is both to find evidence and to reassure readers and reviewers that searches have been thorough. This requires conducting a comprehensive search of all the available sources in order to establish confidence about the inclusiveness of all relevant evidence. Even then, one may be limited by accessibility of evidence.

Observational Studies

During the process of developing the protocol for a systematic review, systematic reviewers need to make decisions as to what study designs are appropriate for answering their research question(s). Based on the nature of the question, outcome, or methodologic preferences, some reviews may include only studies of experimental design (e.g., randomized and/or nonrandomized controlled trials); other reviews may require the addition of observational studies, for example when examining harms outcomes.

By design, RCTs are always hypothesis testing and are considered "confirmatory" studies: they are designed to test the null hypothesis of no difference between the compared groups for a given outcome. Nonrandomized controlled trials and observational studies may be either confirmatory (i.e., hypothesis-testing) or exploratory (i.e., hypothesis-generating) in nature. However, based on a publication alone, it is often difficult to distinguish between confirmatory and exploratory studies. There may be more concern about data dredging in exploratory studies, and the risk of ORB and ARB may be greater than for confirmatory studies.[37]

- We do not recommend searching for registry information for observational studies, as their study registration is not yet mandated and registration is infrequent.
- Reviewers may limit their search for protocols to specific study designs such as trials and prospective observational studies
- We recommend against routinely searching for protocols of retrospective studies. As with RCTs, systematic reviewers can consider contacting study authors for additional information when practical.
- Searching the World Wide Web may be considered as a last option to find protocols of nonrandomized and observational studies.

Grey Literature

The IOM describes grey literature as including trial registries, conference abstracts, books, dissertations, monographs, and reports held by the Food and Drug Administration (FDA) and other government agencies, academics, business, and industry, and standard 3.2.1 recommends that those conducting a systematic review should "search grey literature databases, clinical trial registries, and other sources of unpublished information about studies."[8]

Study Registries

- Reviewers should always search ClinicalTrials.gov and the ICTRP for trials that began recruitment after 2005.
- Match trials with publications found from the standard search, noting 1) trials with an entry in ClinicalTrials.gov, and 2) trials for which no publication was found. ClinicalTrials.gov may provide useful information on results for trials registered after 2008.
- Construct a table that provides information on trials found in the registry, their publication status, and whether they are completed or currently active trials, and provide a count of the number of unique trials found along with their status at the time of the search.

Because of its broader coverage, and because that coverage includes trials registered in ClinicalTrials.gov, we recommend that EPCs always consider conducting a search of the ICTRP in addition to ClinicalTrials.gov. Unpublished studies should be identified by matching studies found in the registry search with publications found in the literature search. This is specifically true for trials that began recruitment after 2008 and for which at least one of the participating centers was based in the United States. Data available from ClinicalTrials.gov is described in Appendix B, while the data from the ICTRP is described in Appendix A.

Regulatory Documents

- Reviewers should search Drugs@FDA for information on drugs; if a search is not conducted reviewers should provide a rationale explaining why the search was not considered necessary or appropriate.
- When reviewers search for evidence at Drugs@FDA, they should focus their search on the Medical Review and Statistical Review documents.

Reviewers should always conduct a search of the FDA CDER Drugs@FDA Website (http://www.accessdata.fda.gov/scripts/cder/drugsatfda/) for Medical and Statistical Reviews relevant to the review. When a search is not conducted, the review should provide a rationale for why the authors believed that a search was not necessary. As an example, consider a comparative effectiveness review (CER) on treatment for migraine. Such a review may require consideration of as many as 20 different drug classes. In such a situation a review of FDA documents may, at present, prove impractical because of the challenges of using the FDA site. In these instance reviewers may choose not to search the FDA site, but they should provide a rationale explaining their reason for not doing so and consider factoring in this limitation in their assessment of the risk of reporting bias.

The Drugs@FDA site may be searched by the generic or trade drug name (not drug class) for Statistical and Medical Reviews written by FDA personnel examining information submitted by pharmaceutical companies for drug approval. However, the Web site typically does not have documents related to older drugs and very new drugs. Reviews should be downloaded and hand searched for trials. The CDER site also lists any post-marketing study commitments that are made after the FDA has approved a product for marketing (e.g., studies requiring the sponsor to demonstrate clinical benefit of a product following accelerated approval).[103]

Information contained in these reviews is typically not adequate to assess trial quality. However, information included in the reviews can identify unpublished studies and unpublished data from published studies, and can be used to verify data obtained from published manuscripts of these trials or to supplement the published results. Studies identified in FDA documents should be compared with those found in the published literature and unpublished studies submitted by manufacturers to identify any remaining unpublished studies or relevant study data not previously published. In addition, the results of the trials reported in the FDA documents should be compared with those reported in published reports of the same studies to identify variation in outcome reporting. However, comparing data from the FDA Medical and Statistical Review documents can be challenging because it is not always easy to identify whether a particular FDA report pertains to a given included study, and it is important to avoid double counting study data in an evidence synthesis.

Study Protocols

- Study protocols that are retrieved in the literature search should be routinely used to identify SOR and SAR.

In the absence of a protocol for an included study, and when feasible, reviewers may consult other relevant sources (such as contacting authors, trial registries, industry and regulatory submissions, and other bibliographic databases not previously searched) not previously searched to attempt to obtain either the protocol or related details reported elsewhere. Because contacting authors is of unclear utility, reviewers may reasonably restrict this exercise to a subset of studies (design, year, country, etc.) or impose other restrictions for which data are likely to be obtained.

Conference Abstracts and Proceedings

- Reviewers should routinely consider conducting a search of conference abstracts and proceedings to identify unpublished or unidentified studies.
- Consult the TEP for suggestions on particular conferences to search and search those conferences specifically.
- Search the full conference abstracts of any meeting identified by reading the references of key articles.
- We do not recommend using conference and meeting abstracts for assessing SOR and SAR, given the variable evidence of concordance between conference abstracts and their subsequent full-text publications.

Current guidance[7] stipulates always including search of databases that index meeting reports, such as Conference Papers Index, Scopus, Papers and Proceedings 1st, BIOSIS previews, etc.. That guidance notes that because the yield is often in the hundreds rather than in the thousands it does not add appreciably to the burden of the review. Current guidance also recommends searching the reports of specific conferences if any Technical Expert Panel (TEP) member or other key informant suggests that the topic of a particular meeting or conference is highly relevant to the topic of the report and searching the full conference abstracts of any meeting that is found by reading the references of other relevant articles.[7]

Grants Database

- Searches of grant databases may aid in the assessment of publication bias in the non industry funded research and may be useful in identifying SOR by comparing grant protocols with published results. However, in the absence of evidence, we are uncertain as to whether such a search will help to identify or minimize outcome or analysis reporting bias. Not all agencies may be involved in the conduct or publication of the studies they fund. Given the additional effort required to do so, we do not recommend contacting nonindustry funding agencies for additional information.
- Searches of grants databases, in general, should only be conducted upon suggestions from the TEP or other key informants.
- Since the process of matching to publications is challenging and the yield likely to be low, when grants databases are searched, we recommend conducting a pilot search first.
- After identifying studies from the grants database, search trial registries using the grant number, title, or name of principal investigator.

- Look for publications of funded grants by searching Medline with the grant number or title.

Since this task is time-consuming, we recommend searching grant databases when review authors anticipate a significant yield in the number of eligible studies. Review authors should search trial registries using grant titles and numbers for each study to determine registration status of eligible studies. The process of finding exact publications is manual and time consuming. Therefore review authors may conduct a pilot search in grant databases to estimate potential yield in eligible studies. After all funded studies are identified, review authors can compare grant description or posted protocols with publications to judge publication bias and selective outcome reporting. Review authors should consider the number of funded studies as a true denominator of the publication bias in National Institutes of Health funded research.

Contacting Authors to Identify Unpublished Studies

- Although likely to occur infrequently, authors should be contacted when in the review team's judgment clarification regarding study eligibility, study design, or other aspects of study conducts is essential to the conduct of the CER and may affect conclusions.
- When authors are contacted, we recommend that no more than three attempts at contact be made, each attempt separated by a week, and that this be done consistently for all authors from whom information is being sought.
- When contacting authors, be clear and concise in your request and, when possible, provide a table identifying the specific data being requested.
- If bias is suspected based on the study report, adding this to the correspondence may help with obtaining information.

IOM standard 3.2.2 recommends that authors of systematic reviews "invite researchers to clarify information about study eligibility, study characteristics, and risk of bias." Although not part of a standard search, and likely to occur infrequently, EPCs should contact researchers and invite them to provide necessary information, when in the review team's judgment clarification regarding study eligibility, study design, or other aspects of study conduct is essential to the conduct of the CER and may affect the conclusions of the review. This might be the case, for example, when disaggregated data is available, and is needed to evaluate benefits and/or harms in sub-populations included in the aggregate data.

Contacting study authors can be time intensive, with uncertain yield and effects on review conclusions. When trying to contact a study author, there is little guidance as to how many times this should be attempted. We were unable to locate any papers providing guidance concerning this point, although a survey (n=111 respondents) of systematic reviewers conducted by Mullan, et al.[79] reported that most respondents contacted at least one study author. Anecdotal experience suggests trying to contact study authors up to three times separated by a week interval between each attempt. To avoid potential bias it seems sensible to make a similar number of contacts with all study authors from whom additional information is sought. Trying to contact one study author three times and other study authors once is systematically different and might introduce bias. We are unaware of any reports examining the possible biases associated with contacting or not contacting study authors. Theoretically, a bias might arise if efforts to contact study authors were systematically different. For example, if the review team were examining the comparative effectiveness of two drug eluting devices and ended up only contacting authors of papers that

systematically provided nonsignificant effect estimates. Therefore, reviewers should consider the possible biasing effects of strategies for contacting study authors and strive to avoid them when possible.

For specific data, such as a missing standard deviation, the review team may want to provide a brief table depicting the missing information. Whatever information is being requested of study authors it is important that the request is made clearly and concisely. It is may be useful to let the study authors know that their help will be acknowledged in the review's report and any subsequent publication.

Contacting Study Sponsors

- When available, EPCs should use industry documents in tandem with published study results for their assessments of risk of ORB and ARB.
- The SRC, rather than EPC staff, should be responsible for contacting primary study sponsors.
- The search for industry documents should include information requested directly from manufactures, as well as industry documents available from the Drug Industry Document Archive.

IOM Standard 3.2.3 states that, in addition to contacting study authors and researchers, authors of systematic reviews should "[i]nvite all study sponsors and researchers to submit unpublished data, including unreported outcomes, for possible inclusion in the systematic review." The request to manufacturers for product information, including information about published and unpublished studies is part of the standard search conducted by the SRC on behalf of the EPCs, and is described in the chapter on Finding Evidence.[7] Industry documents made public as a result of litigation may also be available from the Drug Industry Document Archive (DIDA). When the review team is aware of litigation regarding a drug under review, they should search DIDA for potentially relevant documents.

Handsearching

- Reviewers should routinely conduct a search of the Cochrane Central Register of Controlled Trials.
- If reviewers decide that more comprehensive hand searching is warranted, before conducting the search, work with content experts to identify appropriate journals for hand searching and with a librarian to determine how well those journals are indexed in electronic databases.
- Even when a decision is made not to conduct a general hand search, consider conducting a handsearch of selected key journals to test the sensitivity of the electronic database searches.

IOM Standard 3.2.4 states that authors of systematic reviews should "[h]andsearch selected journals and conference abstracts." Reviewers should routinely conduct a search of the Cochrane Central Register of Controlled Trials (Central), since Central is supplemented with studies gleaned from a hand search of more than 2,000 poorly indexed journals. The Master List, available at http://us.cochrane.org/master-list catalogs the journals and conference abstracts being searched by various Cochrane groups. In addition to routinely searching Central, reviewers should consider on a case-by-case basis whether to conduct handsearches of selected key

journals that are highly relevant to the topic of the report, but not fully indexed, or indexed at all, in the major bibliographic databases, to check the sensitivity of electronic database searches. If the hand search does not identify any relevant studies (or only identifies small and/or lower-quality studies that are unlikely to affect the conclusions of the review) more comprehensive handsearching may be unnecessary. If the reviewers determine that more comprehensive handsearching is necessary, either based on the topic of the systematic review or based on finding missed studies in a selective check of journals, we suggest that they work with content experts to determine which journals may be candidates for handsearches, and with a research librarian to determine which of those journals to hand search, based on how well the journal is indexed in electronic databases and the lag time to indexing.

Information From Searches of the World Wide Web

- We do not recommend that review authors search the World Wide Web for additional information beyond those sources discussed above, unless there are specific reasons to do so
- If the World Wide Web is used as an information source, the rationale for doing so must be clearly presented, along with the methods for searching.

IOM standard 3.2.5 states those conducing systematic reviews should "[c]onduct a web search." Current guidance recommends using Web of Science or Scopus if they are available. If subscriptions to these services are not available, however, current guidance recommends using Google Scholar rather than other free search engines such as PubReMiner or PubFocus.[7] However, given the lack of evidence, we are uncertain of the utility of searching the World Wide Web to locate additional information on a given study and do not recommend including such a search as part of the standard or expanded search for evidence unless there is a compelling reason to do so. When a Web search is conducted, a clear rationale for doing so should be presented, along with specific information about the nature of the search, as well as a description of what was retrieved and how that information was screened and included information selected.

Searching for Non-English Language Literature

- Reviewers should avoid the use of English-language only filters when searching standard databases.
- Abstracts and other reports of non-English language studies should be tracked to inform a judgment of the likelihood of bias that might arise from excluding non-English language reports.
- Discuss with the TEP whether excluding non-English language articles might bias the findings of the report.
- Search databases that specifically index reports of studies in languages other than English 1) when a review of English-language abstracts suggests systematic differences between studies reported in English language journals and those reported in non-English language journals, or 2) based on information from TEP members or other key informants.

IOM standard 3.2.6 states that those conducting systematic reviews should search for studies reported in languages other than English if appropriate. Searches of databases that specifically index non-English language literature, however, are likely to be the exception, rather than the rule. On the other hand, a review of English language abstracts of non-English language articles,

retrieved during the standard search of the major bibliographic databases, can inform the decision regarding the need for a more comprehensive search for non-English language articles. This is why current guidance recommends against the use of English-only filters when searching major bibliographic databases.[7] If a comparison of the English-language abstracts of non-English articles finds consistent systematic differences in results with articles published in English, the review team should consider expanding the search to include non-English language articles. In addition, the review team should discuss with the TEP whether exclusion of non-English studies might bias the report. When an assessment based on these criteria suggests that non-English language articles be included, we recommend a staged approach. Such an approach might initially include a further review of all English language abstracts of non-English language articles found as part of the standard search. Findings from this review might then suggest expanding the search to include special regional databases.

The review team should always review the English language abstracts retrieved in the search of the major bibliographic databases. The literature search should be expanded to include databases that specifically index non-English language literature such as LILACs (Literatura Latino Americana e do Caribe em Ciências da Saúde) and Global Index Medicus when a review of the abstracts finds:

1. A consistent difference between studies reported in English-language abstracts of non-English language studies and those reported in other languages;
2. Most of the relevant studies have been reported in a language other than English; or
3. Most of the studies have been conducted in non-English language regions.

Identification of Selective Outcome Reporting and Selective Analysis Reporting based on the Study Report

- Efforts should routinely be made to identify outcome level SOR and SAR for each study included in a systematic review.
- In general, systematic reviewers should recognize that studies that do not investigate or report outcomes of interest to the review may be susceptible to SOR or SAR, and so should not exclude studies such studies from the review.
- We suggest restricting outcome and analysis reporting bias assessments to those outcomes that will be graded for their strength of evidence.
- Collate all companion publications for a given study.
- Compare the outcomes and analyses specified in the methods section to those presented in the results section, looking for discrepancies.

SOR and SAR may be identified from information contained within the published report. The systematic reviewer should thus start with the study report(s) to try and identify SOR and SAR. The first step is to collect all companion reports related to the study of interest that provide information on the study methodology and that report outcomes of interest to the systematic reviewer. Online appendices and other data referenced in the reports should always be collected and examined. Next, comparisons should be made between the methods and the results section of the study report(s), looking specifically for discrepancies that may represent one of the types of SOR and SAR listed in Table 1.

Limitations to the Proposed Approach

There are limitations to relying on the study publication for identifying SOR and SAR. In particular, discrepancies between the methods and results sections cannot be reliably considered as adequate assessment of ORB and ARB because manuscripts are prepared at a late stage in the research process, generally after authors have reviewed the results and decided which data will be presented. As such, the Methods section of the report may already have been selectively tailored to support favorable findings. It should be noted, however, that our assessments of ORB and ARB specifically reflect a *risk* as it applies to our review, as opposed to actual bias in the study (which may or may not be present). For example, authors may be genuinely limited by journal word count restrictions and hence report some outcomes in narrative form or completely omit reporting them altogether.

Process for Identifying Selective Outcome Reporting and Selective Analysis Reporting and for Assessing Outcome Reporting Bias and Analysis Reporting Bias

This section outlines the process for assessing the risk of ORB and ARB, incorporating the guidance and procedures outlined above. In general, systematic reviewers should not exclude studies that do not investigate or report outcomes of interest to the review. These studies still need to be assessed for the risk of ORB and ARB.

Above we described and categorized the various types of reporting biases in published studies. We also reported evidence-based recommendations for both detecting and minimizing ORB and ARB. For example, data gleaned from trial registries might inform salient aspects of the trial protocol and lead to detection of outcome or analysis reporting bias affecting the published evidence. When previously unreported outcomes data are also obtained from a registry, the potential for bias would be mitigated for that outcome.

For each gradable outcome of benefit or harm, reviewers will need to assign an appropriate risk of reporting bias to their included studies irrespective of whether they did or did not contribute outcome data. While there is no best or universally accepted approach, we suggest one option that EPCs might consider. The approach is as yet untested. Alternatively, EPCs might devise other systematic approaches that should be reported in the Methods section of the review.

As with all steps and tasks in a systematic review, the information gained and its potential effects on the conclusions of a review must be balanced with the resources (time, effort, and expertise) needed to accomplish the task. Nevertheless, a routine and systematic search for information from additional sources including study protocols, registries, regulatory submissions, industry documents, and authors is necessary before one can have a sense of what might be necessary and to reassure readers of the review that you have been thorough.

Proposed Steps

Assessment of selective reporting bias for a study is outcome specific. We recommend assessment only of outcomes that are to be graded for their strength of evidence. For a given systematic review, study outcomes data are at no risk of reporting bias if all the gradable outcomes that inform a systematic review are fully reported, even if others were concealed. In this case, no further action is needed.

While assessing selective reporting bias, we recommend that all companion reports of a study should be linked and examined together. A two step decision aid is suggested (Figure 1). In step

1, a preliminary assessment of the risk of ORB and ARB is made. Step 2 involves clarifying or modifying the preliminary risk assessments by evaluating additional study information from sources other than included study report(s). Step 2 also involves making judgments about the likelihood or risk of outcome reporting bias given the aims and objectives, design, duration, and measurement of other outcomes.

Assessing the Risk of Outcome Reporting Bias and Analysis Reporting Bias: Step 1

During this step, the systematic reviewer compares the Methods sections with the Results sections for a given outcome (denoted outcome 'X' in the following scenarios). If a published protocol for that study could be accessed during the routine systematic literature search, comparing the protocol with the Results section will be more informative for determining the risk of ORB or ARB. During this comparison process, refer to Table 1 for identifying the types of SOR and SAR.

Study Risk of ORB and ARB Ruled Out – Scenario 1

When outcome X was reported either in the Methods section or the protocol (when available through regular searches), or both, and completely reported in the study report(s), and appropriately analyzed as planned, then the study is not at risk for ORB or ARB ("ORB risk –" or "ARB risk –"). No further assessment is necessary and the risk of ORB and ARB may be determined at this stage.

Study at Risk of ORB or ARB – Scenario 2

If outcome X was planned (i.e., specified in the Methods section or the protocol) but the results were not reported or were partially reported in the Results section, then the study is at risk of ORB for that outcome ("ORB risk +"). Also, when reported results are based on a different analysis, effect measure, cut-off, etc. than what was prespecified, then the study is at risk of ARB for that outcome ("ARB risk +").

Study at Risk of ORB or ARB – Scenario 3

If outcome X was not described in the Methods section or protocol (when available), but the results were reported or incompletely reported, then the study is at risk of ORB for that outcome ("ORB risk +"). This study is also at risk of ARB because there is no way to know whether the reported analysis was planned or *post hoc*.

Study Risk of ORB and ARB Could Not Be Ruled Out – Scenario 4

If outcome X is not reported in the Methods or Results sections of the study publication(s) and no other information is available, risk of SOR and SAR cannot be ruled out ("unclear risk of ORB"). This would also apply to a study that did not report any outcome of review interest but was eligible on population, intervention, comparator, and other criteria. Similarly, if the analyses for outcome X are not specified in the Methods section, then the risk of ARB for that outcome cannot be ruled out ("unclear risk of ARB").

Summary of Scenarios 2 Through 4

In scenarios 2 to 4, if a study was deemed "ORB/ARB risk +" or "unclear risk of ORB/ARB", then further investigation (where feasible and appropriate) is required to clarify the risk. This occurs at Step 2.

Assessing the Risk of Outcome Reporting Bias and Analysis Reporting Bias: Step 2

In Step 2 the reviewer attempts to strengthen or modify the initial assessment of ORB and ARB with additional information obtained from an external source document (e.g., FDA report, trial registry data, etc.). For a given outcome, the additional information could serve to identify discrepancies between planned analyses and results (thereby providing new information to inform selective reporting bias risk), provide outcome data that were previously missing or inappropriately reported in the included study literature (thereby modifying our preliminary assessment of bias for that outcome), or both. For this step, however, EPCs may decide whether cross checking against external source documents is feasible or relevant based on the guidance reported above for each potential source; if not, this needs to be documented in the systematic review.

As in Step 1, when additional data are available from unpublished or external sources (such as protocols provided by Investigators, FDA data, etc.), the reviewer should reassess the risk of ORB and ARB and document the exact source(s) of information used for that assessment.

For studies for which the outcomes are compared with an external source document and no discrepancies in planned and reported outcomes and analyses are identified, or additional outcome data that were previously missing or inappropriately reported in the included study literature are obtained, these can be deemed as studies for which risk of ORB and ARB could not be detected (i.e., ORB and/or ARB risk negative). If a study is judged to be affected by ORB and ARB (i.e., ORB and/or ARB risk positive), or the risk of ARB could not be ruled out because no additional documents were located (unclear risk of ARB), then the assessment may be deemed completed. In situations when additional data are not available or do not inform a preliminary assessment of "unclear risk of ORB", a final assessment described below is recommended.

Final Assessment for Outcome Reporting Bias Risk That Cannot be Ruled Out

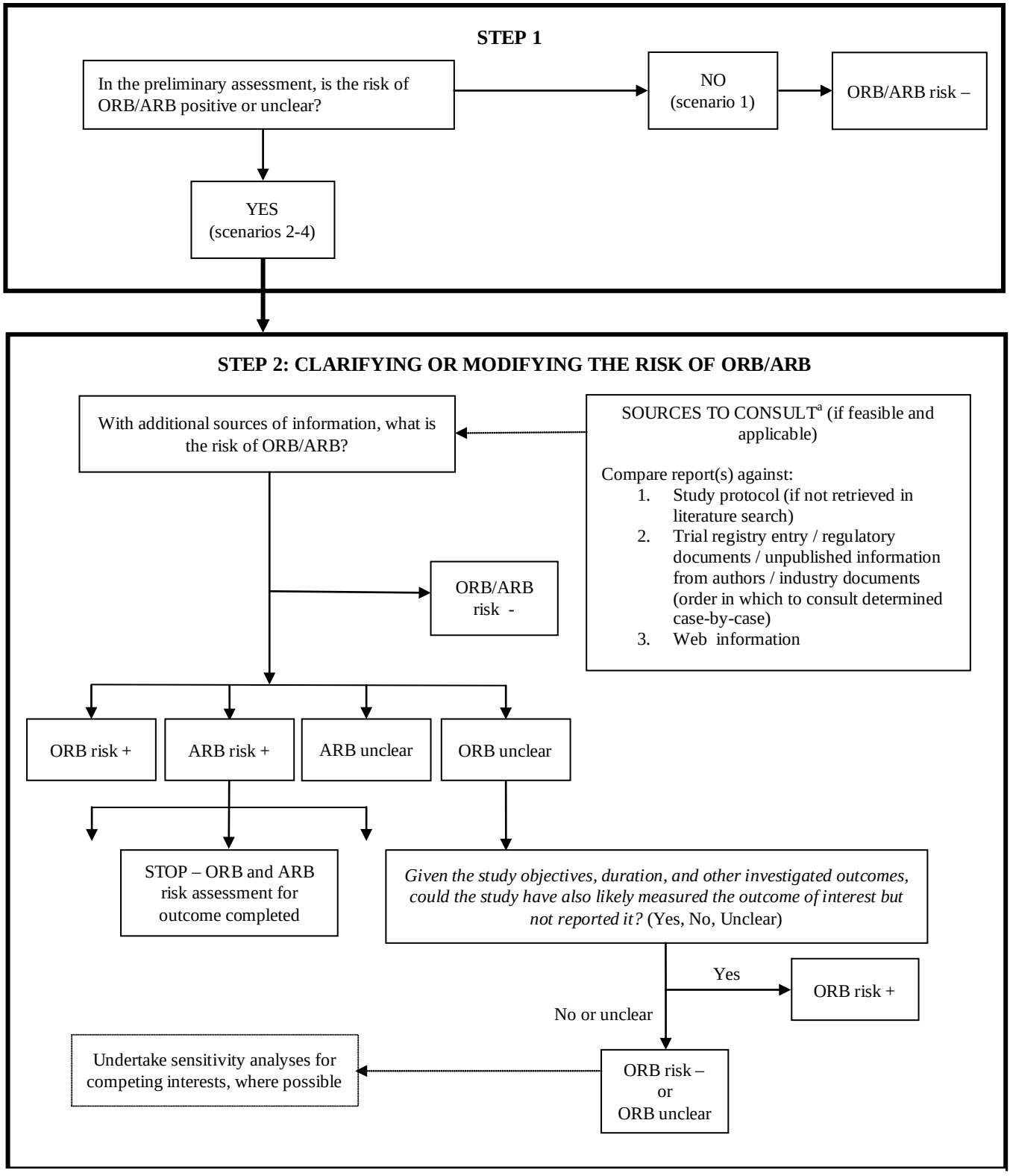
For studies for which the risk of ORB cannot be ruled out, we suggest that EPCs do one final assessment. Reviewers should ask the question: "*Given the study objectives, duration, and other investigated outcomes, could the study have also likely measured the outcome of interest but not reported it?*" If the answer is "no" the study should be rated as "ORB risk negative". If it still remains unclear whether the outcome of interest may have been assessed, the study should be categorized as "ORB risk unclear." Alternatively, when the answer is "yes" (e.g., another reported outcome in the study leads the reviewer to believe that outcome X would have been collected), then the study should be rated "ORB risk positive" for that outcome. This should be done for all included studies for all gradable outcomes, not just those that reported outcomes data. As such it is important that systematic reviewers should not exclude studies that do not investigate or report outcomes of interest to the review without a sound rationale.

Alternatively, EPCs could also construct a matrix as described by Dwan, et al.[104] using a multi-step process that reviewers can use to determine if potentially eligible trial reports are prone to ORB:

- Include all included studies (accompanied with all corresponding publications) irrespective of whether they report the review-relevant outcomes. Unless justified otherwise, studies should not have been excluded because they did not report any of the review outcomes.

- Document, for all included studies, which ones reported all, at least one but not all, or none of the review-relevant outcomes.
- Identify included studies that did not report one or more review-relevant outcomes.
- For included studies that did not report or only partially reported one or more review-relevant outcomes, construct a matrix table including the review-relevant outcomes as well as those reported in these studies, by arranging outcomes in columns and studies in rows. This matrix should differentiate complete, partial, and no reporting for each outcome.

Figure 1. Flow diagram of the risk of outcome reporting bias and analysis reporting bias assessment process



Note: ARB=analysis reporting bias, ORB=outcome reporting bias.
[a] Document exact source of information that clarifies or modifies concern of ORB or ARB.

Risk of Bias Assessment for Selective Outcome Reporting and Selective Analysis Reporting

We recommend that individual study risk of ORB and ARB should be categorized as ORB/ARB risk positive, unclear, or risk negative. This assessment for individual studies will inform an EPC's overall assessment of the risk of reporting bias as recommended in the strength of evidence guidance.

Comparisons with Other Guidance

Kirkham, et al.[23] developed a classification termed the Outcomes Reporting Bias in Trials (ORBIT) system. The ORBIT designation of SOR is based only on information contained within the study publication(s) and from clinical judgment. The ORBIT approach to SOR provides a useful paradigm for thinking about SOR. ORBIT researchers advocate developing a matrix of review outcomes reported in each included study, using information from within the study publications as well as clinical judgment as to what outcomes should have been reported in studies on a particular topic.

We do not recommend the routine use of the ORBIT classification system for SOR for several reasons. First, the scope and intended purpose of ORBIT does not encompass SOR that occurs when researchers add outcomes that were not prespecified. In addition, SOR assessment from ORBIT relies on information contained within the study as well as clinical judgment, but does not incorporate information obtained from sources such as trial registries or study protocols.

We are unaware of other classification systems for the types of SOR and SAR.

Combining When Publication Bias or Outcome Reporting Bias is Suspected

The decision regarding whether to combine studies and how to report the result necessarily depends on the level of suspicion of bias. In some cases, the best course is to refrain from combining the available studies if it is known that an amount of data that could influence results is being withheld. For example, the manufacturer Pfizer initially refused to provide data for all of its reboxetine trials for an Institute for Quality and Efficiency in Health Care (IQWiG) review.[105, 106] Since data on only about 1600 out of 4600 patients were analyzed, IQWiG concluded that no statement of benefit or harm could be made. After negative publicity, Pfizer provided the data, and the subsequent IQWiG review reported no benefit of reboxetine for depression.

The funnel plot is a scatter plot of precision versus treatment effect, with a point for each study. The plot is interpreted visually with asymmetric appearance suggesting (presumably negative) studies that may not have been published. Statistical methods based on funnel plot have been proposed to detect and adjust for publication bias. However, for assessing publication bias, an international group of methodologists has concluded that funnel plot has very limited application in meta-analysis.[107] Sensitivity analyses can assess whether a finding of treatment benefit is robust to differing assumptions regarding the extent of potential bias.[108-110] However, empirical validation of sensitivity analyses has not been possible, because the true extent of bias in any particular review is unknown. Furthermore, sensitivity methods do not help pin down the size of the effect, which varies depending on the amount of bias assumed.

When there is no avenue for discovering hidden studies and no applicable statistical method for assessing publication bias, sensitivity analyses should be considered and the potential for bias should be noted when reporting combined data.

Reporting the Search Strategy and Results

General Guidelines

As described more fully in the chapter on Finding Evidence,[7] reviews should provide complete strategies for all indexed databases that were included in the search. Strategies should be included in the appendices of AHRQ publications, and authors should offer to include them as part of the supplementary material offered online for any journal publications. In addition, to the items described in Finding Evidence, the following information should be reported:

- If trial registries or regulatory documents are searched, a count of unpublished studies identified through the trial registries or regulatory documents should be reported.
- If authors or primary studies are contacted, the review should report the authors contacted and the associated study, the number of attempted contacts, and whether the contact was successful.
- Reports of hand searches should include the journals searched and how they were selected, and potentially relevant citations should be recorded and tracked for inclusion in the PRISMA diagram.
- In general, whenever recommended guidelines are not followed, the review should include a rationale for that decision.

Reporting of Findings and Investigations of Reporting Bias

Systematic reviews must provide the reader with transparent and reproducible methods and results in regards to efforts to identify the risk of ORB and ARB. Each review requires a thoughtful, individualized approach to identifying SOR and SAR, which must be outlined in the review, along with the rationale for that approach. Most importantly, the rationale for decisions to explore, or to not explore, information sources outside of the study publication should be clearly presented to the reader.

Some recommendations for avoiding and addressing outcome reporting bias can be gleaned from a tutorial on the assessment of completed reviews.[104] A matrix of trials by outcomes reported can be constructed. When this is done, trials should not be excluded because they do not report, or only partially report, outcomes of interest. Instead, evidence that the missing outcomes were measured should be noted, as well as the level of suspicion that suppression was related to the results. Refraining from reporting summary estimates should be reserved for cases with a high level of suspicion of the deliberate withholding of a substantial proportion of data. Although empirical validation of sensitivity analyses has not been possible, a combination of cautious reporting and sensitivity analyses is preferable in cases where there is potential selective reporting. At a minimum, we suggest that the following steps should be described in a systematic review (in evidence tables) for included studies:

- For each gradable outcome, reviewers should report their final study ORB/ARB risk assessments similar to their reporting of study risk of bias assessments by outcomes.
- Include the citation to the study protocol with the citations for the main study publications.
- If additional information from a trial protocol, registry, or regulatory submission documents was used to assess SOR or SAR, describe what that specific information was and how it contributed to the identification of SOR/SAR, and the assessment of ORB and ARB.

- To help readers assess the extent of outcome reporting bias, systematic reviewers should cross-tabulate trials versus reported outcomes.
- For each included study, reviewers should report the study funder or sponsor and the conflicts of interest of the study authors.
- In reviews where the existence of unobtainable studies has been verified, reviewers should express their opinion concerning the risk of publication bias.

References

1. Stevens A. Madame de Staël: a study of her life and times, the first revolution and the first empire: Harper & Brothers; 1881.

2. Gibson L. GlaxoSmithKline to publish clinical trials after US lawsuit. BMJ. 2004 Jun 26;328(7455):1513. PMID: 15217848.

3. Turner EH, Matthews AM, Linardatos E, et al. Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy. N Engl J Med. 2008;358(3):252-60.

4. Melander H, Ahlqvist-Rastad J, Meijer G, et al. Evidence b(i)ased medicine--selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. BMJ. 2003 May 31;326(7400):1171-3. PMID: 12775615.

5. Whittington CJ, Kendall T, Fonagy P, et al. Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. Lancet. 2004 Apr 24;363(9418):1341-5. PMID: 15110490.

6. Song F, Parekh S, Hooper L, et al. Dissemination and publication of research findings: an updated review of related biases. Health Technol Assess. 2010;14(8):1-193.

7. Relevo R, Balshem H. Finding evidence for comparing medical interventions: Agency for Healthcare Research and Quality (AHRQ) and the Effective Health Care program. Journal of Clinical Epidemiology. 2011 Jun 16PMID: 21684115.

8. Committee on Standards for Systematic Reviews of Comparative Effectiveness Research, Institute of Medicine. Finding What Works in Health Care: Standards for Systematic Reviews: Natl Academy Pr; 2011.

9. Dickersin K, Chalmers I. Recognising, investigating and dealing with incomplete and biased reporting of clinical research: from Francis Bacon to the World Health Organisation. 2010. http://www.jameslindlibrary.org/essays/biased_reporting/biased_reporting.html.

10. Sterling TD. Publication Decisions and Their Possible Effects on Inferences Drawn from Tests of Significance--Or Vice Versa. Journal of the American Statistical Association. 1959;54(285):30-4.

11. Sterling TD, Rosenbaum WL, Weinkam JJ. Publication decisions revisited: the effect of the outcome of statistical tests on the decision to publish and vice versa. American Statistician. 1995:108-12.

12. Dickersin K. The Existence of Publication Bias and Risk Factors for Its Occurrence. JAMA. 1990;263(10):1385-9.

13. Dwan K, Altman DG, Arnaiz JA, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. PLoS One. 2008;3(8):e3081.

14. McAuley L, Pham B, Tugwell P, et al. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? Lancet. 2000 Oct 7;356(9237):1228-31. PMID: 11072941.

15. Olson CM, Rennie D, Cook D, et al. Publication bias in editorial decision making. JAMA. 2002 Jun 5;287(21):2825-8. PMID: 12038924.

16. Pham B, Platt R, McAuley L, et al. Is there a "best" way to detect and minimize publication bias? An empirical evaluation. Eval Health Prof. 2001 Jun;24(2):109-25. PMID: 11523382.

17. Rising K, Bacchetti P, Bero L. Reporting Bias in Drug Trials Submitted to the Food and Drug Administration: Review of Publication and Presentation. PLoS Med. 2008;5(11):e217.

18. Veitch E. Tackling Publication Bias in Clinical Trial Reporting. PLoS Medicine. 2005;2(10).

19. Sterne JAC, Egger M, Moher D. Chapter 10: Addressing reporting biases. In: Higgins JTP, Green S, eds. Cochrane handbook for systematic reviews of interventions. Version 5.1.0 [updated March 2011] ed.: The Cochrane Collaboration; 2011.

20. Chan A-W, Hrobjartsson A, Haahr MT, et al. Empirical Evidence for Selective Reporting of Outcomes in Randomized Trials: Comparison of Protocols to Published Articles. JAMA. 2004;291(20):2457-65.

21. Dwan K, Altman DG, Cresswell L, et al. Comparison of protocols and registry entries to published reports for randomised controlled trials. Cochrane Database Syst Rev. 2011(1):MR000031. PMID: 21249714.

22. Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses. BMJ. 2012;344:d7202. PMID: 22214754.

23. Kirkham JJ, Dwan KM, Altman DG, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. BMJ. 2010;340:c365. PMID: 20156912.

24. Kirkham JJ, Altman DG, Williamson PR. Bias due to changes in specified outcomes during the systematic review process. PLoS One. 2010;5(3):e9810. PMID: 20339557.

25. Zarin D. Newsmaker interview: Debora Zarin. Unseen world of clinical trials emerges from U.S. database. Interview by Eliot Marshall. Science. 2011 Jul 8;333(6039):145. PMID: 21737714.

26. De Angelis C, Drazen JM, Frizelle FA, et al. Clinical Trial Registration: A Statement from the International Committee of Medical Journal Editors. Annals of Internal Medicine. 2004;141(6):477-8.

27. Zarin DA, Tse T, Williams RJ, et al. The ClinicalTrials.gov results database--update and key issues. New England Journal of Medicine. 2011 Mar 3;364(9):852-60. PMID: 21366476.

28. Higgins JTP, Green S, eds. Cochrane handbook for systematic reviews of interventions. Version 5.1.0 [updated March 2011] ed: The Cochrane Collaboration; 2011.

29. Higgins JTP, Altman DG, Sterne JAC. Chapter 8: Assessing risk of bias in included studies. In: Higgins JTP, Green S, eds. Cochrane handbook for systematic reviews of interventions. Version 5.1.0 [updated March 2011] ed.: The Cochrane Collaboration; 2011.

30. Food and Drug Administration Modernization Act of 1997. 21 USC 301. U.S.A; 1997.

31. International Committee of Medical Journal Editors. ICMJE: Uniform Requirements for Manuscripts Submitted to Biomedical Journals. 2010. http://www.icmje.org/urm_main.html. Accessed on November 19, 2010.

32. Food and Drug Administration Amendments Act of 2007. 21 USC 301. U.S.A; 2007.

33. Law MR, Kawasumi Y, Morgan SG. Despite Law, Fewer Than One In Eight Completed Studies Of Drugs And Biologics Are Reported On Time On ClinicalTrials.gov. Health Aff (Millwood). 2011 Dec;30(12):2338-45. PMID: 22147862.

34. National Library of Medicine. About ClinicalTrials.gov - ClinicalTrials.gov. 2012. http://clinicaltrials.gov/ct2/info/about. Accessed on May 10, 2012.

35. World Health Organization. International Clinical Trials Registry Platform (ICTRP). 2012. http://apps.who.int/trialsearch/AdvSearch.aspx. Accessed on May 10, 2012.

36. Williams RJ, Tse T, Harlan WR, et al. Registration of observational studies: is it time? CMAJ. 2010 Oct 19;182(15):1638-42. PMID: 20643833.

37. Loder E, Groves T, Macauley D. Registration of observational studies. BMJ. 2010;340:c950. PMID: 20167643.

38. The Lancet. Should protocols for observational research be registered? The Lancet. 2010;375(9712):348-.

39. Pearce N. Registration of protocols for observational research is unnecessary and would do more harm than good. Occup Environ Med. 2011 Feb;68(2):86-8. PMID: 21118848.

40. Sorensen HT, Rothman KJ. The prognosis for research. BMJ. 2010;340:c703. PMID: 20164129.

41. Wood AJJ. Progress and Deficiencies in the Registration of Clinical Trials. New England Journal of Medicine. 2009;360(8):824-30.

42. Prayle AP, Hurley MN, Smyth AR. Compliance with mandatory reporting of clinical trial results on ClinicalTrials.gov: cross sectional study. BMJ. 2012;344:d7373. PMID: 22214756.

43. Viergever RF, Ghersi D. The quality of registration of clinical trials. PLoS One. 2011;6(2):e14701. PMID: 21383991.

44. Shamliyan T. Reporting of results of interventional studies by the information service of the National Institutes of Health. Clinical Pharmacology: Advances and Applications. 2010;2:169-76.

45. Reveiz L, Chan AW, Krleza-Jeric K, et al. Reporting of methodologic information on trial registries for quality assessment: a study of trial records retrieved from the WHO search portal. PLoS One. 2010;5(8):e12484. PMID: 20824212.

46. Ross JS, Tse T, Zarin DA, et al. Publication of NIH funded trials registered in ClinicalTrials.gov: cross sectional analysis. BMJ. 2012;344:d7292. PMID: 22214755.

47. Ross JS, Mulvey GK, Hines EM, et al. Trial Publication after Registration in ClinicalTrials.Gov: A Cross-Sectional Analysis. PLoS Medicine. 2009;6(9):e1000144.

48. Wieseler B, Kerekes MF, Vervoelgyi V, et al. Impact of document type on reporting quality of clinical drug trials: a comparison of registry reports, clinical study reports, and journal publications. BMJ. 2012;344:d8141. PMID: 22214759.

49. Golder S, Loke YK, Bland M. Unpublished data can be of value in systematic reviews of adverse effects: methodological overview. Journal of Clinical Epidemiology. 2010 May 8;63(10):1071-81.

50. Mathieu S, Boutron I, Moher D, et al. Comparison of registered and published primary outcomes in randomized controlled trials. JAMA. 2009 Sep 2;302(9):977-84. PMID: 19724045.

51. Rising K, Bacchetti P, Bero L. Correction: Reporting Bias in Drug Trials Submitted to the Food and Drug Administration: Review of Publication and Presentation. PLoS Medicine. 2009 January 01, 2009;6(1):e17.

52. Scharf O, Colevas AD. Adverse event reporting in publications compared with sponsor database for cancer clinical trials. J Clin Oncol. 2006 Aug 20;24(24):3933-8. PMID: 16921045.

53. Sweet BV, Schwemm AK, Parsons DM. Review of the processes for FDA oversight of drugs, medical devices, and combination products. J Manag Care Pharm. 2011 Jan-Feb;17(1):40-50. PMID: 21204589.

54. Man-Son-Hing M, Wells G, Lau A. Quinine for nocturnal leg cramps: a meta-analysis including unpublished data. J Gen Intern Med. 1998 Sep;13(9):600-6. PMID: 9754515.

55. Floyd JS, Serebruany VL. Prasugrel as a potential cancer promoter: review of the unpublished data. Arch Intern Med. 2010 Jun 28;170(12):1078-80. PMID: 20585076.

56. Nissen SE, Wolski K, Topol EJ. Effect of muraglitazar on death and major adverse cardiovascular events in patients with type 2 diabetes mellitus. JAMA. 2005 Nov 23;294(20):2581-6. PMID: 16239637.

57. Natanson C, Kern SJ, Lurie P, et al. Cell-free hemoglobin-based blood substitutes and risk of myocardial infarction and death: a meta-analysis. JAMA. 2008 May 21;299(19):2304-12. PMID: 18443023.

58. Ofman JJ, MacLean CH, Straus WL, et al. A metaanalysis of severe upper gastrointestinal complications of nonsteroidal antiinflammatory drugs. J Rheumatol. 2002 Apr;29(4):804-12. PMID: 11950025.

59. Jefferson T, Jones MA, Doshi P, et al. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. Cochrane Database Syst Rev. 2012;1:CD008965. PMID: 22258996.

60. O'Connor AB. The need for improved access to FDA reviews. JAMA. 2009 Jul 8;302(2):191-3. PMID: 19584349.

61. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline: Guideline for good clinical practice E6(R1) International Conference on Harmonisation. June 10, 1996.

62. Horton R. Pardonable revisions and protocol reviews. Lancet. 1997 Jan 4;349(9044):6. PMID: 8988113.

63. McNamee D, James A, Kleinert S. Protocol review at The Lancet. Lancet. 2008 Jul 19;372(9634):189-90. PMID: 18640443.

64. Godlee F. Publishing study protocols: making them visible will improve registration, reporting and recruitment. BMC News and Views. 2001;2:4-.

65. Altman DG, Furberg CD, Grimshaw JM, et al. Lead editorial: trials - using the opportunities of electronic publishing to improve the reporting of randomised trials. Trials. 2006;7:6. PMID: 16556322.

66. Chan A-W, Krleza-Jeric K, Schmid I, et al. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. Canadian Medical Association Journal. 2004;171(7):735-40.

67. Hahn S, Williamson PR, Hutton JL. Investigation of within-study selective reporting in clinical research: follow-up of applications submitted to a local research ethics committee. J Eval Clin Pract. 2002 Aug;8(3):353-9. PMID: 12164983.

68. Vedula SS, Bero L, Scherer RW, et al. Outcome reporting in industry-sponsored trials of gabapentin for off-label use. N Engl J Med. 2009 Nov 12;361(20):1963-71. PMID: 19907043.

69. Hartling L, Bond K, Vandermeer B, et al. Applying the risk of bias tool in a systematic review of combination long-acting beta-agonists and inhaled corticosteroids for persistent asthma. PLoS One. 2011;6(2):e17242. PMID: 21390219.

70. Royen PV, Sandholzer H, Griffiths F, et al. Are presentations of abstracts at EGPRN meetings followed by publication? European Journal of General Practice. 2010;16(2):100-5.

71. Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. Cochrane Database Syst Rev. 2007(2):MR000005. PMID: 17443628.

72. Tam VC, Hotte SJ. Consistency of phase III clinical trial abstracts presented at an annual meeting of the American Society of Clinical Oncology compared with their subsequent full-text publications. Journal of Clinical Oncology. 2008 May 1;26(13):2205-11. PMID: 18445846.

73. Chokkalingam A, Scherer R, Dickersin K. Agreement of data in abstracts compared to full publications. Controlled Clinical Trials. 1998;19(3):S61–S2-S–S2.

74. Hopewell S, Clarke M, Askie L. Reporting of trials presented in conference abstracts needs to be improved. Journal of Clinical Epidemiology. 2006;59(7):681-4.

75. Toma M, McAlister FA, Bialy L, et al. Transition from meeting abstract to full-length journal article for randomized controlled trials. JAMA. 2006;295(11):1281-.

76. Shamliyan T, Kane RL. Clinical research involving children: registration, completeness, and publication. Pediatrics. 2012 May;129(5):e1291-300. PMID: 22529271.

77. Shamliyan TA, Kane RL, Wyman J, et al. Results availability from clinical research of female urinary incontinence. Neurourology and Urodynamics. 2011.

78. Shamliyan TA, Kane RL, Wyman J, et al. Results availability from clinical research of female urinary incontinence. Neurourol Urodyn. 2012 Jan;31(1):22-9. PMID: 22038753.

79. Mullan RJ, Flynn DN, Carlberg B, et al. Systematic reviewers commonly contact study authors but do so with limited rigor. Journal of Clinical Epidemiology. 2008 Nov 13.

80. Young T, Hopewell S. Methods for obtaining unpublished data. Cochrane Database of Systematic Reviews. John Wiley & Sons, Ltd; 1996.

81. Kyzas PA, Loizou KT, Ioannidis JP. Selective reporting biases in cancer prognostic factor studies. J Natl Cancer Inst. 2005 Jul 20;97(14):1043-55. PMID: 16030302.

82. U.S. Cochrane Center. Resources for handsearchers | US Cochrane Center. 2011. http://us.cochrane.org/resources-handsearchers. Accessed on April 21, 2012.

83. Hopewell S, Clarke M, Lefebvre C, et al. Handsearching versus electronic searching to identify reports of randomized trials (Review). Cochrane Database of Systematic Reviews. 2007 Apr 18(2).

84. Glanville J, Cikalo M, Crawford F, et al. Handsearching did not yield additional unique FDG-PET diagnostic test accuracy studies compared with electronic searches: a preliminary investigation. Research Synthesis Methods. 2012.

85. Armstrong R, Jackson N, Doyle J, et al. It's in your hands: the value of handsearching in conducting systematic reviews of public health interventions. J Public Health (Oxf). 2005 Dec;27(4):388-91. PMID: 16311247.

86. Egger M, Zellweger-Zähner T, Schneider M, et al. Language bias in randomised controlled trials published in English and German. The Lancet. 1997;350(9074):326-9.

87. Heres S, Wagenpfeil S, Hamann J, et al. Language bias in neuroscience--is the Tower of Babel located in Germany? Eur Psychiatry. 2004 Jun;19(4):230-2. PMID: 15196606.

88. Vickers A, Goyal N, Harland R, et al. Do Certain Countries Produce Only Positive Results? A Systematic Review of Controlled Trials. Controlled Clinical Trials. 1998;19(2):159-66.

89. Egger M, Juni P, Bartlett C, et al. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. Health technology assessment (Winchester, England). 2003;7(1):1-76.

90. Gregoire G, Derderian F, Le Lorier J. Selecting the language of the publications included in a meta-analysis: is there a Tower of Babel bias? J Clin Epidemiol. 1995 Jan;48(1):159-63. PMID: 7853041.

91. Jüni P, Holenstein F, Sterne J, et al. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. International Journal of Epidemiology. 2002;31(1):115-23.

92. Moher D, Pham B, Klassen TP, et al. What contributions do languages other than English make on the results of meta-analyses? J Clin Epidemiol. 2000 Sep;53(9):964-72. PMID: 11004423.

93. Morrison A, Moulton K, Clark M, et al. English-language restriction when conducting systematic review-based meta-analyses: systematic review of published studies. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2009. p. 1-17.

94. Pham B, Klassen TP, Lawson ML, et al. Language of publication restrictions in systematic reviews gave different results depending on whether the intervention was conventional or complementary. Journal of Clinical Epidemiology. 2005;58(8):769-76.e2--76.e2.

95. Schulz KF, Chalmers I, Hayes RJ, et al. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA. 1995 Feb 1;273(5):408-12. PMID: 7823387.

96. Li J, Burnham JF, Lemley T, et al. Citation Analysis: Comparison of Web of Science®, Scopus™, SciFinder®, and Google Scholar. Journal of Electronic Resources in Medical Libraries. 2010;7(3):196-217.

97. Kulkarni AV, Aziz B, Shams I, et al. Comparisons of citations in Web of Science, Scopus, and Google Scholar for articles published in general medical journals. JAMA. 2009 Sep 9;302(10):1092-6. PMID: 19738094.

98. Bakkalbasi N, Bauer K, Glover J, et al. Three options for citation tracking: Google Scholar, Scopus and Web of Science. Biomed Digit Libr. 2006;3:7. PMID: 16805916.

99. Jasco P. As we may search—Comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced databases. Current Science. 2005;89(9):1537-47.

100. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions--agency for healthcare research and quality and the effective health-care program. J Clin Epidemiol. 2010 May;63(5):513-23. PMID: 19595577.

101. Gotzsche PC, Jorgensen AW. Opening up data at the European Medicines Agency. BMJ. 2011;342:d2686. PMID: 21558364.

102. Lim E, Brown A, Helmy A, et al. Composite outcomes in cardiovascular research: a survey of randomized trials. Ann Intern Med. 2008 Nov 4;149(9):612-7. PMID: 18981486.

103. Drug Effectiveness Review Project. Systematic Review Methods and Procedures Oregon Health & Science University. Portland, Oregon: Revised January 2011. http://www.ohsu.edu/xd/research/centers-institutes/evidence-based-policy-center/derp/documents/upload/DERP_METHODS_WEB_Final_January-2011-2.pdf

104. Dwan K, Gamble C, Kolamunnage-Dona R, et al. Assessing the potential for outcome reporting bias in a review: a tutorial. Trials. 2010;11:52. PMID: 20462436.

105. McGauran N, Wieseler B, Kreis J, et al. Reporting bias in medical research - a narrative review. Trials. 2010;11(1):37.

106. Wieseler B, McGauran N, Kaiser T. Finding studies on reboxetine: a tale of hide and seek. BMJ. 2010;341:c4942. PMID: 20940211.

107. Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011;343:d4002. PMID: 21784880.

108. Copas JB, Shi JQ. A sensitivity analysis for publication bias in systematic reviews. Stat Methods Med Res. 2001 Aug;10(4):251-65. PMID: 11491412.

109. Vevea JL, Woods CM. Publication bias in research synthesis: sensitivity analysis using a priori weight functions. Psychol Methods. 2005 Dec;10(4):428-43. PMID: 16392998.

110. Copas J, Jackson D. A bound for publication bias based on the fraction of unpublished studies. Biometrics. 2004 Mar;60(1):146-53. PMID: 15032784.

Abbreviations

Abbreviation	Definition
AHRQ	Agency for Healthcare Research and Quality
ARB	Analysis reporting bias
CADTH	Canadian Agency for Drugs & Technology in Health
CDER	FDA Center for Drug Evaluation and Research
CDRH	FDA Center for Devices and Radiological Health
CONSORT	CONsolidated Standards of Reporting Trials
EMA	European Medicines Agency
EPC	Evidence-based Practice Center
FDA	U.S. Food and Drug Administration
FDAAA	U.S. Food and Drug Administration Amendments Act
FDAMA	U.S. Food and Drug Administration Modernization Act
GSK	GlaxoSmithKline
ICMJE	International Committee of Medical Journal Editors
ICTRP	International Clinical Trials Registry Platform
MeSH	Medical Subject Headings
NCT	National Clinical Trial number
NDA	New Drug Application
ORB	Outcome reporting bias
ORBIT	Outcomes Reporting Bias in Trials
PMA	Premarket Application
RCT	Randomized controlled trial
RePORT	Federal Research Portfolio Online Reporting Tools
SAE	Serious adverse event
SAR	Selective analysis reporting
SIP	Scientific information packet
SOR	Selective outcome reporting
SRC	Scientific Resource Center
TEP	Technical Expert Panel
WAME	World Association of Medical Editors
WHO	World Health Organization