

***Methods Guide
for Comparative Effectiveness Reviews***

Finding Grey Literature Evidence and Assessing for Outcome and Analysis Reporting Biases When Comparing Medical Interventions: AHRQ and the Effective Health Care Program



Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

This report is based on research led by the Oregon and Ottawa Evidence-based Practice Centers (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract Nos. 290-2007-10057-I and 290-2007-10059-I). 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.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

This document was written with support from the Effective Health Care Program at AHRQ. This document is in the public domain and may be used and reprinted without special permission. Citation of the source is appreciated.

Persons using assistive technology may not be able to fully access information in this report. For assistance contact info@ahrq.gov.

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

Suggested citation: Balshem H, Stevens A, Ansari M, Norris S, Kansagara D, Shamliyan T, Chou R, Chung M, Moher D, Dickersin K. Finding Grey Literature Evidence and Assessing for Outcome and Analysis Reporting Biases When Comparing Medical Interventions: AHRQ and the Effective Health Care Program. Methods Guide for Comparative Effectiveness Reviews. (Prepared by the Oregon Health and Science University and the University of Ottawa Evidence-based Practice Centers under Contract Nos. 290-2007-10057-I and 290-2007-10059-I.) AHRQ Publication No. 13(14)-EHC096-EF. Rockville, MD: Agency for Healthcare Research and Quality. November 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Authors:

Howard Balshem, M.S.^a
Adrienne Stevens, M.Sc.^b
Mohammed Ansari, M.D., M.Med.Sc., M.Phil.^b
Susan Norris, M.D., MPH, M.Sc.^a
Devan Kansagara, M.D.^c
Tatyana Shamliyan, M.D., M.S.^d
Roger Chou, M.D., MPH^a
Mei Chung, Ph.D, MPH^c
David Moher, Ph.D.^b
Kay Dickersin, Ph.D.^f

The authors are all members of the AHRQ EPC Reporting Bias Workgroup.
Following are additional members of the workgroup:

Donna M. Dryden, Ph.D.
Jennifer Lin, M.D.
Rose Relevo, MLIS, M.S.
Alexander Tsertsvadze, M.D., M.Sc.

- a. Oregon Health & Science University, Portland, OR
- b. Ottawa Hospital Research Institute, Ottawa, Ontario, Canada
- c. Portland VA Medical Center, Portland, OR
- d. University of Minnesota School of Public Health, Minneapolis, MN; Elsevier Evidence Based Medicine Center, Philadelphia, PA
- e. Tufts Medical Center, Boston, MA
- f. Johns Hopkins Bloomberg School of Public Health, Baltimore, MA

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.

Strong methodological approaches to systematic review improve the transparency, consistency, and scientific rigor of these reports. Through a collaborative effort of the Effective Health Care (EHC) Program, the Agency for Healthcare Research and Quality (AHRQ), the EHC Program Scientific Resource Center, and the AHRQ Evidence-based Practice Centers have developed a Methods Guide for Comparative Effectiveness Reviews. This Guide presents issues key to the development of Systematic Reviews and describes recommended approaches for addressing difficult, frequently encountered methodological issues.

The Methods Guide for Comparative Effectiveness Reviews is a living document, and will be updated as further empiric evidence develops and our understanding of better methods improves. We welcome comments on this Methods Guide paper. 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 email to epc@ahrq.hhs.gov.

Richard G. Kronick, Ph.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 and Task Order Officer
Evidence-based Practice Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Peer Reviewers

Peter Doshi
Johns Hopkins University
Baltimore, MD

Kerry Dwan
University of Liverpool
Liverpool, United Kingdom

Eileen Erinoff
ECRI Institute
Philadelphia, PA

Jonathan Leo
Lincoln Memorial University
Harrogate, TN

Yoon Loke
University of East Anglia
Norwich, United Kingdom

Andreas Lundh
Nordic Cochrane Center
Copenhagen, Denmark

Natalie McGauren
Institute for Quality and Efficiency in Health Care
Cologne, Germany

Key Points

- Reviews of the literature consistently provide evidence of significant reporting biases.
- Reporting bias should be cautiously assumed to exist even if authors cannot determine its direction and magnitude. As such, all included studies must be assessed for reporting bias.
- When studies do not investigate or report outcomes of interest to the review this may be due to a reporting bias.
- Assessment of outcome and analysis reporting bias should be restricted to those outcomes that will be graded for their strength of evidence, for feasibility.
- Sources of Evidence
 - Reviewers should always search ClinicalTrials.gov and the International Clinical Trials Registry Platform
 - Reviewers should routinely search and request clinical study reports from the European Medical Agency, and should search Drugs@FDA for Medical and Statistical Review documents
 - Study protocols should be sought during the literature searching process
 - Reviewers should routinely consider searching conference abstracts and proceedings to identify unpublished or unidentified studies and should consult with their Technical Expert Panels for specific conferences to search
 - Reviewers should routinely conduct a search of the Cochrane Central Register of Controlled Trials, a source of handsearching results
 - Reviewers should avoid the use of English-only filters when searching standard databases
 - Searches of grant and non-English databases and contact with authors may be warranted
 - The utility of these sources for identifying or minimizing reporting bias associated with observational studies has not yet been evaluated.
- All sources of evidence, with the exception of conference abstracts, should be collated and used for assessing selective outcome and analysis reporting biases. A framework for assessing selective reporting is detailed. If reviewers decide to use the framework for observational studies, certain considerations or adaptations of the framework may need to be made.

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.^{2,3} 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.⁴

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 may be 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 had once thought. For this reason, we are recommending searching these other sources for studies that might otherwise not be 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. As 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. As a third example, given the fact that only 60 percent of randomized controlled trials (RCTs) described in conference abstracts reach full publication,⁵ and full publication is associated with results favoring the test intervention, then conference abstracts from the meeting(s) most likely to publish trial abstracts are probably worth searching. That said, before conducting their own search, the systematic

reviewers should check sources such as the Cochrane Collaboration's Central Register of Controlled Trials to make sure this task hasn't already been done by others.

The earlier guidance chapter by Relevo and Balslem⁶ (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. Nor do we discuss the problem of ghostwriting, which is a question of appropriately and transparently attributing authorship. 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 14 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 of 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 as well as unpublished data from published studies. The second workgroup focused on how to identify and assess the likelihood of biases arising from selective outcome and selective analysis reporting. Each workgroup member participated in one or more subgroups. While we considered techniques for assessing the likelihood of publication bias outside the scope of this guidance, some approaches for assessing publication bias were addressed by the second workgroup.

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 500 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 these EPC internal reviewers 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 was revised again based on peer review and public comment. However, the final guidance reflects the views of the authors and the EPC program, and not those of the peer or public reviewers.

This guidance is divided into four parts. The first part provides an introduction to the guidance, describes the methodology used to develop the guidance, and provides some brief background information on reporting bias. Part 2 describes the major sources of evidence that can be used to minimize the risk of missing information relevant to the review, discusses the available evidence on the value of searching each source, and provides recommended guidance on using each source. Part 3 provides guidance on the process of assessing for selective reporting of outcomes and analyses. Finally, Part 4 offers brief guidance on reporting the search strategy and results.

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 health care interventions.”⁷ 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.⁸ 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.⁹ He confirmed his findings 35 years later in a second survey,¹⁰ and to this day new research continues to demonstrate the existence of sizable publication bias.¹¹⁻¹⁸ Box 1 describes several types of reporting biases that have been identified in the literature.

Box 1. Definitions of some types of reporting biases^a

Publication bias

The publication or nonpublication of research because of the nature and direction of the results.

Time lag bias

The rapid or delayed publication of research because of the nature and direction of the results.

Multiple publication bias

The multiple or singular publication of research because of the nature and direction of the results.

Location bias

The publication of research in journals with different ease of access or levels of indexing in standard databases because of the nature and direction of results.

Citation bias

The citation or noncitation of research because of the nature and direction of the results.

Language bias

The publication of research in a particular language because of the nature and direction of the results.

Outcome reporting bias

The selective reporting, in published studies, of one or more outcomes because of the nature and direction of the results.

Analysis reporting bias

The selective reporting, in published studies, of one or more analyses as a change from planned analyses or as a selection from two or more analysis options because of the nature and direction of the results.

^aAdapted from definitions provided in the Cochrane Handbook for Systematic Reviews of Interventions.¹⁹

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 and analysis reporting bias occur when outcomes are selectively reported or data selectively analyzed—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 the evaluation of 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 from 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.

The selective reporting of outcomes and analyses 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 Several studies provide empirical evidence of the effect of selective outcome reporting and selective analysis reporting on the pooled estimates of treatment effects.²²⁻²⁵ In addition, the selective reporting of analyses and outcomes may also operate at the systematic review level.^{21-23, 26-29}

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, investigators may choose outcomes based on whether they will produce favorable results, rather than on their importance for clinical practice or policy decision making. Given the aims, objectives, and duration of a study, a strong suspicion in the minds of reviewers that a key outcome of interest was excluded from the study results, which most investigators would not have excluded, should in itself be taken as a signal for risk of selective outcome reporting bias, despite good agreement between study results reporting and study protocol. In other words, the failure to address clinically important outcomes may introduce a form of outcome reporting bias, if studies with negative results for that outcome are less likely to be published. During results analysis, bias occurs if investigators decide to change their analysis (e.g., change in time point) in order to present favorable results or report the most favorable of the several analyses undertaken. 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 the types of selective outcome reporting and selective analysis reporting that could be identified and determined when assessing studies. Some of these constructs are also listed elsewhere.^{30, 31}

Table 1. Types of selective outcome and analysis reporting, which may affect the direction and/or magnitude of the reported study findings

Selective Outcome Reporting	Selective Analysis Reporting
<p>Missing/changed outcomes:</p> <ul style="list-style-type: none"> • Omission of an outcome that was prespecified or for which the clinical judgment of the review team strongly suggests should have been prespecified • Addition of an outcome that was not prespecified (excluding unintended or unanticipated harms outcomes) • Change from the protocol in a primary or secondary outcome • Failure to report prespecified subgroups • Reporting of a composite outcome without reporting of results for individual components, or reporting of composites of unconventional components • Use of a different outcome measurement tool or definition from that prespecified in the protocol without a reasonable justification • Incomplete specification of an outcome domain (e.g., ‘substance use’ versus ‘abstinence’ or ‘reduction in use’) and specific measurement (e.g., self-reported measures versus levels in biologic tissues) in the methods section of the publication or in other available sources <p>Incomplete reporting</p> <ul style="list-style-type: none"> • Partial reporting of outcomes (in other words, information is not sufficient to add the study to a meta-analysis) for example: including an 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”) 	<p>Changes to/in (planned), or selection from (multiple):</p> <ul style="list-style-type: none"> • Data types, for example, dichotomous instead of continuous using favorable post hoc cut-offs • Effect measure specific metric or method of aggregation, for example, reporting of the more favorable of the change-from-baseline (change score) or the final value comparison for a continuous outcome when both were analyzed • Assumptions of data distribution or estimate adjustments without reasonable justification • Time points for analysis • Post hoc subgroup 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.”⁷ This section discusses

the various sources of data discussed in the IOM report, provides empirical evidence of their value as sources of information both for unpublished studies and for unpublished data in published studies, as well as evidence that excluding evidence from these sources can lead to biased effect estimates, and recommends how these sources can be used in the search for evidence.

Grey Literature

The IOM describes grey literature as including trial registries, conference abstracts, books, dissertations, monographs, and reports held by the U.S. Food and Drug Administration (FDA) and other government agencies, academics, business, and industry. 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.”⁷ Our recommendations for incorporating grey literature in the guidance below apply specifically to reviews of conventional drugs and devices (Table 2).

Table 2. Recommended sources of grey literature for conventional drugs and devices

Type of Information	Recommended Sources or Strategies	When To Search	Reporting Bias Type	Provisos	Recommendation(s) Empiric Evidence (E) or Consensus (C)
Study protocol elements (Methods), outcomes data (Results), or completely missing studies	ClinicalTrials.gov ICTRP	Routinely	SOR/SAR/ Publication bias	Studies conducted in 2005 and onwards	E
Study protocol elements (Methods), outcomes data (Results), or completely missing studies	FDA EMA	Routinely	SOR/SAR/ Publication bias	Indication approved drugs, and class III devices	E
Missing studies	Conference abstracts and proceedings	Routinely, on advice of KI or TEP	Publication bias		E
Missing studies	Grant databases (e.g. Research Portfolio Online Reporting Tools)	On KI or TEP recommendation	Publication bias		C
Study protocol elements (Methods) or outcomes data (Results)	Study Authors	For data clarifications (regarding study eligibility, study design, or other aspects of study conducts)	SOR/SAR	No more than three attempts	C

Table 3. Recommended sources of grey literature for conventional drugs and devices (continued)

Type of Information	Recommended Sources or Strategies	When To Search	Reporting Bias Type	Provisos	Recommendation(s) Empiric Evidence (E) or Consensus (C)
Study protocol elements (Methods), outcomes data (Results), or completely missing studies	Industry SIPs, Industry maintained trial registries, and DIDA	Routinely for SIPs and DIDA At reviewers discretion for Industry maintained trial registries	SOR/SAR/publication bias	EPCs should not contact the Industry directly, SIPs through SRC	C
Study protocols, companion papers, or completely missing studies	Hand searching	Routinely search the Cochrane Central Register of Controlled Trials Hand searching of selected journals at reviewers discretion	SOR/SAR/Location bias		C
Study protocols, companion papers, or completely missing studies	Non-English language literature	Search routinely ^a	Language bias		C
Study protocols, companion papers, or completely missing studies	Citation searching using the World Wide Web	Not recommended	SOR/SAR/Publication bias		C

Note: DIDA = Drug Industry Document Archive, EMA = European Medical Agency, EPC = Evidence-based Practice Center, FDA = U.S. Food and Drug Administration, ICTRP = International Clinical Trials Registry Platform, KI = Key Informant, SAR = selective analysis reporting, SIP = scientific information packet, SOR = selective outcome reporting, SRC=Scientific Resource Center, TEP = Technical Expert Panel.

^aSearch criteria only, not eligibility criteria. If non-English language literature is excluded, a list of potentially relevant but excluded literature can help inform the potential risk of language bias.

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 FDA Modernization Act of 1997³² 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.²⁸ The FDA Amendments Act of 2007³³ further required 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 (see Table 3). However, it's important to note that the FDA Amendments Act does not cover trials initiated and completed before 2007, and so will not cover older drugs unless they are tested in trials that were either initiated or ongoing in 2007.³⁴ ClinicalTrials.gov, launched in 2000 to comply with FDA Modernization Act, currently contains over 139,000 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 182 countries.³⁵ 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.³⁶ Appendix B describes the data elements available from the ICTRP.

Observational studies, where the assignment of subjects into a treated group versus a control group is outside the control of the investigator, 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 2010.³⁷ 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,³⁷⁻³⁹ 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.^{40, 41}

Table 4. Registration and reporting requirements of the U.S. Food and Drug Administration Amendments Act, Section 801^a (reprinted with permission from Wood 2009⁴²)

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	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	- 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

Note: 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 late⁴³ and only 12 to 22 percent of trials posted results within one year of completion.^{43, 44} The number of unregistered trials and those with missing results is unknown, as is the accuracy of the data submitted.²⁷ Compliance with the FDA Amendments Act mandatory reporting requirement of trial results is low: within one year of study completion, only 22 percent of 738 trials were compliant.⁴⁴ In a review of a sample of trials registered with the ICTRP between June 2008 and June 2009, Viergever and Ghersi⁴⁵ 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.⁴⁶ 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 ClinicalTrials.gov and six other registries supported by the WHO search portal ICTRP.⁴⁷

In a study of National Institutes of Health funded trials registered in ClinicalTrials.gov, Ross, et al.⁴⁸ 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.⁴⁹ 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, et al. 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.⁵⁰

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, 46, 49, 51-54} 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, 46, 49, 51, 52, 54}

Guidance on Using 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 existing publication, and (2) trials for which no publication was found.
- 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. However, because ICTRP does not require results reporting, systematic reviewers will always want to directly search 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. While mandatory

reporting of results in ClinicalTrials.gov came into effect in Dec 2007, the registry was launched in 2000. The ICMJE required prospective trial registration as a precondition for publication in 2005. This latter date coincides with the launch of ICTRP and appears a reasonable cut-off for when the registries should be searched.

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 are quite different.⁵⁵ 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 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 Medical 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. Finally, the FDA officer reports may also provide authors of systematic reviews with a list of potential studies for inclusion that may not have been found through other sources.

Drugs@FDA, (<http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm>) 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).

Device Approval Process

Medical devices are regulated by the FDA Center for Devices and Radiological Health, 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.⁵⁵ 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-eluting 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 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 Premarket Application process.

While not as useful as Drugs@FDA, a Web-based, searchable database of information about FDA-approved devices (Devices@FDA) is available at <http://www.accessdata.fda.gov/scripts/cdrh/devicesatfda/index.cfm>.

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;⁵⁶ that FDA documents suggested an elevated risk of harms not acknowledged in FDA advisory committee recommendations;^{57,58} that prompt analysis of data available to the FDA can identify harms not identified in the published literature;⁵⁹ and that publication is associated with positive outcomes;²⁵ but also found 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.⁶⁰ 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.⁶¹

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.⁶² 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. Turner, et al., when comparing the results of unpublished trials of second-generation antipsychotics found in FDA documents with the results of published trials, found that the effect size of the unpublished trials was significantly less than half that for the published trials.²⁴

Data from the FDA may be obtained by searching the FDA Web site, submitting a Freedom of Information Act request, or both. Over a period of several weeks to months the FDA releases the data in the form of their medical and statistical reviews. However, even when available, FDA reviews can be difficult to 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.⁶³

New policies of the European Medical Agency allow access to regulatory submissions with minimal, commercially sensitive redaction, and will soon provide access to clinical trial data for medications it considers for approval.⁶⁴⁻⁶⁶ A review of documents released under the 2010 policy providing access to all documents held by the Agency, suggests that the European Medical Agency should be considered a valuable and routine source of regulatory documents on drug studies.⁶⁴

Guidance on Using Regulatory Documents

- Reviewers should routinely search and request Clinical Study Reports from the European Medical Agency.
- 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 routinely search and request Clinical Study Reports from the European Medical Agency. Reviewers should also conduct a search of the FDA CDER Drugs@FDA Web site (www.accessdata.fda.gov/scripts/cder/drugsatfda/) for Medical and Statistical Reviews relevant to the review and consider submitting Freedom of Information Act requests for drug and class III device trial data early in the course of their systematic review to allow for FDA response time, which could be several weeks. When a search of these sources 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 this 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).⁶⁷

Information contained in these reviews may not be 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

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.^{68, 69}

Study protocols and related information can be located and accessed from several sources such as study authors, industry registries, trial registries, Web sites of relevant agencies (e.g., ClinicalTrials.gov, canadatrials.com, controlled-trials.com, and WHO ICTRP), and through documents made public as a result of litigation. Also, several peer reviewed medical journals including *The Lancet*, *Trials*, and others publish study protocols, or summaries of protocols with full protocols available upon request. *The Lancet* began publishing protocols of randomized trials in 1997 and extended this to observational studies in 2001.^{70, 71} BioMed Central began publishing protocols for a variety of study designs in 2001.⁷² In 2006, the journal *Trials* was launched and has accepted study protocols from the outset.⁷³

Empirical Findings on the Value of Searching for Protocols

Several empirical studies comparing protocols and published reports of individual trials for consistency and completeness of outcomes and analyses^{20, 54, 74, 75} 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, 54, 74, 75} 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.³ 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 on the country in which the trial was conducted, 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.⁷⁶

Guidance on Searching for Study Protocols

- For a priori study methods, grey literature may be a helpful source in the absence of access to full protocol.
- Study protocols that are retrieved in the literature search should be routinely used to

identify selective outcome and analysis reporting.

When the protocol for an included study is not found as part of the standard search, reviewers should include other relevant sources such as contacting authors and searching trial registries, industry sites, regulatory submissions, and bibliographic databases not previously searched to attempt to obtain either the protocol or protocol-related details. Since protocols are frequently amended, reviewers should search for later amendments and cross validate the currency of study protocols against Clinical Study Reports submitted to regulatory agencies and using the “history” function of ClinicalTrials.gov.

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, Van 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).⁷⁷ 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.⁵ Tam and Hotte⁷⁸ 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.^{79–81}

Guidance on Using 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 selective outcome reporting and selective analysis reporting, given the variable evidence of concordance between conference abstracts and their subsequent full-text publications.

Abstract and conference proceedings should be searched as a source for identifying trials that may not otherwise be published or which might have been missed in the initial search.

Current guidance⁶ stipulates always including search of databases that index meeting reports, such as Conference Papers Index, Scopus, Papers and Proceedings 1st, BIOSIS previews, et cetera. 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.⁶

Grant Databases

Several grant databases allow for analysis of the 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 grants 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 Web site 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.^{82, 83} 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.^{82, 83} The analysis of NIH funded pediatric trials demonstrated that only 33 percent were registered and only 53 percent were published.⁸² The analysis of NIH funded therapeutic studies for female urinary incontinence found that only 6 percent were registered.⁸³ Published studies (94 percent of all NIH funded) mentioned the NIH grant numbers but did not necessarily report study results.⁸³

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.

Guidance on Using Grants Databases

- 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 and Key Informants or the TEP 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.

Contacting Authors

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 only for the whole study population 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 and 93 were published in high-impact 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.⁸⁴ Missing data was the most common reason for contacting study authors.

In a systematic review of the literature on methods for obtaining unpublished data, Young, et al. 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.⁸⁵

Three studies not considered in the Young review assessed whether contacting authors for more information adds substantive information. Kyzas, et al.⁸⁶ 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). Young, et al. noted, however, that response rates do not seem to be influenced by the number of requests.⁸⁵

Chan, et al.⁷⁴ 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. An additional limitation to contacting authors is that they may not have access to full data, or may be contractually obligated to nondisclosure

Guidance on Contacting Authors

- 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 conduct 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.
- When reviewers contact authors, they should report the number of authors they attempted to contact, the number of authors actually contacted, and the percentage of authors who responded positively to the request for 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 only disaggregated data are reported, and there is a need 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. An additional limitation to contacting authors is that they may not have access to full data or may be contractually obligated to nondisclosure. 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.⁸⁴ 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 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

Some pharmaceutical companies have started to publicly share their own trial registry data. GlaxoSmithKline has announced that it will release all anonymized patient level data since 2007 in their Clinical Study Register (www.gsk-clinicalstudyregister.com).^{87, 88} Novo Nordisk also provides Web access to its trial registry.⁸⁹ EPC literature searches for published studies are routinely supplemented with a request to the manufacturer for a scientific information packet (SIP). The SIP includes information about products available from the product label as well as information about published and unpublished trials or studies about the product. To ensure consistency in the way SIPs are requested and to ensure transparency by eliminating contact between the EPC conducting the review and the manufacturers of products being reviewed, the Scientific Resource Center for the AHRQ Effective Health Care Program routinely requests SIPs from manufacturers on behalf of the EPCs for all CERs and technical briefs.

Empirical Findings on the Value of Contacting Study Sponsors

Limited evidence exists on the use of industry documents for identifying selective outcome and analysis reporting, and has been mainly obtained through legal proceedings. Vedula et al. compared 12 of 20 internal pharmaceutical company documents with their published versions (1999-2006) for off-label use of gabapentin.³ The authors found discrepancies in the primary outcome in the publications of 8 of 12 trials (new primary outcome, no distinction between primary and secondary outcomes, change from primary to secondary outcomes, or outcomes omitted), with statistically significant results presented in five publications. Psaty and Kronmal compared mortality data of two published trials with their respective internal pharmaceutical company documents for rofecoxib given for Alzheimer disease or cognitive impairment² In both publications, mortality data were provided in narrative form without accompanying statistical analyses, whereas statistically significant hazard ratios were reported in the internal documents.

Jefferson, et al. recounted their unsuccessful experience trying to obtain unpublished data on oseltamivir from the manufacturer and recommended requesting the full clinical study reports for each trial, but noted there is no guarantee those reports are reliable.⁹⁰

Guidance on Contacting Study Sponsors

- When available, EPCs should use industry documents in tandem with published study results for their assessments of risk of outcome and analysis reporting biases.
- The SRC, rather than EPC staff, should be responsible for contacting primary study sponsors for Scientific Information Packets.
- The search for industry documents should include information requested directly from manufactures, as well as industry documents available from the Drug Industry Document Archive.
- Reviewers may also consider searching publicly accessible trial registries maintained by GlaxoSmithKline Inc. and Novo Nordisk Inc.

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 Scientific Resource Center on behalf of the EPCs, and is described in the guidance on Finding Evidence.⁶ 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. Additional sources that may be searched include:

GlaxoSmithKline: www.gsk-clinicalstudyregister.com

Novo Nordisk: www.novonordisk-trials.com/website/content/trial-results.aspx

However, given that there is little data on the completeness, accuracy, or usefulness of industry-maintained trial registries, as well as the lack of evidence that including such data does not tilt the weight of evidence synthesis in favor of one company over another, we hesitate to make a strong recommendation for searching these additional sources of grey literature.

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, process of care topics, 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.⁹¹ 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).⁹² 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 of 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 with database searching.⁹³

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.⁹⁴ 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.

Guidance on 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.

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.

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. Therefore, even when systematic reviewers have not placed language restrictions on searches or inclusion criteria, 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 WHO and LILACS (Latin American and Caribbean Literature in Health Sciences).

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

A MEDLINE search of all publications from 2000 to February 3, 2011, conducted by the author of this section, found that of 6,574,939 citations, 90 percent were published in English. Table 4 shows the number and frequency of publications in other languages with at least 1 percent frequency.

Table 5. 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 the author’s 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 IOM standards for (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,^{95, 96} and almost all Chinese acupuncture trials published in Chinese report positive results.⁹⁷ Numerous other studies, however, have found that excluding non-English language publications may not have an impact on the conclusions in systematic reviews.⁹⁸⁻¹⁰⁴

Guidance on 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.⁶ 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 of non-English language articles 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. Systematic differences between studies reported in English language journals and those reported in non-English language journals;
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.

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. The World Wide Web is a platform for citation-searching databases as opposed to grey literature searches per se.

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.¹⁰⁵⁻¹⁰⁸ These studies found that Web of Science, Scopus, and Google Scholar produced quantitatively and qualitatively different citation counts, and that each database missed linking to some references included in other databases. None of these studies provided strong evidence that routinely searching the Web has an important impact on review findings.

Guidance on Conducting 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 conducting 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.⁶ However, given the lack of evidence, we are uncertain of the utility of searching the World Wide Web to locate additional studies 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. Because there is no strong evidence showing that routinely searching the Web would have an important impact on review findings, and because of the significant resource burden 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.

Guidance on the Process of Assessing for Selective Reporting of Outcomes and Analyses

This section explains how the risk of outcome and analysis reporting biases can be assessed and clarified once information on a study has been retrieved. The proposed assessments of outcome and analysis reporting biases 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 assessing reporting biases 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 reporting bias 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 reporting bias risk assessments.

Principles for Assessing 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. Reporting bias 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 or masking 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

For outcomes of interest to the review, we suggest restricting reporting bias assessments to those outcomes that will be graded for their strength of evidence according to guidance provided by the EPC Program.¹⁰⁹ Gradable outcomes are those determined a priori during the topic refinement phase and reported in the protocol to be important for health care 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.^{51,110} Reporting biases related to 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 prespecified as an outcome to be included in the assessment of the strength of evidence. In this situation, a post hoc decision may then be made to assess the risk of reporting bias specifically for that outcome.

Composite Outcomes

Reporting of composite outcomes, without reporting on component outcomes, may be a signal of reporting bias.¹¹¹ 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.¹¹² Studies that report composite outcomes should also provide results for the component outcomes.

Reviewers should be suspicious when unexpected components are included or expected components are excluded. For example, in a trial on the effect of hormone replacement therapy on cardiovascular events in recently postmenopausal women, the authors' primary endpoint was a composite outcome of death, admission to hospital for heart failure, and myocardial infarction.¹¹³ Neither stroke nor angina were included, raising concerns whether it was a planned outcome.¹¹⁴

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, protocol data elements reported in clinicaltrials.gov, and methods sections in the articles). In general, systematic reviewers should recognize that when studies do not investigate or report outcomes of interest to the review this may be due to a reporting bias. Missing outcomes should, therefore, not be considered as a criteria for excluding otherwise eligible 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 and Analysis Reporting in Included Studies

Above we described the various sources of information on study outcomes and analyses; the empirical evidence on the accuracy, completeness, and feasibility of using those sources to identify and characterize selective reporting; and guidance on using those sources. In this section, we suggest a procedure for using those sources to assess for reporting biases while conducting a systematic review. Our recommendations apply mostly to experimental studies. For observational studies we provide distinct recommendations. Our recommendations are likely to be revisited as new or more robust evidence emerges.

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 relevant to the objective of the review 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, trials 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. Observational studies may be either confirmatory (i.e., hypothesis-testing) or exploratory (i.e., hypothesis-generating) in nature. Although the risk of selective reporting of the most favorable of multiple analyses exists for both RCTs and observational studies, the risk is much higher when studies are exploratory. 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 reporting biases may be greater than for confirmatory studies.³⁸

Guidance on Including Observational Studies

- 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, observational 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.

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

- As described below, efforts should routinely be made to identify outcome level selective outcome and analysis reporting 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 selective outcome or analysis reporting, and so should not exclude 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 (except conference abstracts) for a given study.
- Compare the planned outcomes and analyses as stated in the Methods section of the report, protocol and other source documents with those reported in the results section, looking for discrepancies.

Comparing Methods Section With Results of Published Reports To Judge the Risk of Outcome Reporting and Analysis Reporting Bias

There are limitations to relying on the study publication for identifying the selective reporting of outcomes or analyses. In particular, discrepancies between the Methods and Results sections cannot be reliably considered as adequate assessment of reporting bias 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 reporting biases 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.

Dichotomization of outcomes data into published and unpublished is overly simplistic. The risk of reporting bias is largely dependent upon the reviewers' access, or lack of access, to all study source documents — peer-reviewed journal reports and their published companion reports, trial registries, abstracts and conference proceedings, regulatory submissions, industry maintained registries and databases, and unpublished data with authors and sponsors. Because selective reporting may not be convincingly identified from information contained within the published study report and its published companion papers, systematic reviewers should endeavor to retrieve as much of the recommended grey literature as possible before undertaking an assessment of the risk of reporting bias.

Proposed Steps

Assessment of selective reporting bias for a study is outcome specific. 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 for reporting bias, we recommend that all companion reports (i.e., published or unpublished data) of a study be linked and examined together (Figure 1). When all the study data from various sources are examined together, concerns about reporting bias provisionally suspected in the study publication might be eliminated because, for example, they were obtained from regulatory submissions or another source of grey literature. On the other hand, reporting bias not otherwise suspected in a trial publication might come to light when compared, for example, with study protocol or trial registry data. Thus, the assessment of reporting bias must be made across all included companion reports—published and retrieved grey literature. EPCs may decide whether cross checking against all recommended external source documents is feasible or relevant based on the guidance reported above for each potential source; if not, this needs to be documented with rationale in the systematic review.

Reviewers should refer to Table 1 for identifying the types of selective reporting impacting the outcome, and categorize their risk assessment as positive, negative or unclear keeping in mind the four levels of measurement specification that have been described by Zarin et al.²⁹ These include

- Domain— e.g. anxiety
- Specific measurement—e.g. Hamilton Anxiety Rating Scale
- Specific metric—e.g. change from baseline at a specified time, and
- Method of aggregation—e.g. categorical with proportion of patients with decrease \geq 50 percent

Following are possible scenarios that may be encountered with respect to a hypothetical outcome X:

Scenario 1—Reporting Bias Ruled Out

When it is clear to the reviewers that outcome X was planned (e.g. from protocol, regulatory submissions, etc.), complete outcome data are available from at least one study document (published or otherwise), and the outcome was appropriately analyzed as planned, then the study is not at risk for reporting bias for this outcome (“ORB risk–” or “ARB risk–”). Here and below

“ORB” and “ARB” refer to “outcome reporting bias” and “analysis reporting bias,” respectively. No further assessment is necessary.

Scenario 2—Clear Risk of Reporting Bias

If reviewers determine that an outcome X was planned but the results were not reported, or were only partially reported in study documents, then the study is at risk of reporting bias 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 analysis reporting bias for that outcome (“ARB risk +”). No further assessment is necessary.

Scenario 3—Clear Risk of Reporting Bias

If reviewers determine that an outcome X was not planned but the results were reported, then the study is at risk of reporting bias for that outcome (“ORB risk +”). This study is also at risk of analysis reporting (“ARB risk +”) because there is no way to know whether the reported analysis was planned or post hoc. No further assessment is necessary.

Scenario 4—Reporting Bias Cannot be Ruled Out

If the reviewers are unable to determine whether an outcome X was planned, but data are reported completely or partially, then the study risk of outcome and analysis reporting bias may be categorized as “unclear”. This would also apply to a study that did not report any outcome of review interest across all source documents but was eligible on population, intervention, comparator, and other criteria. Whenever reviewers have categorized their assessment as “unclear risk of ORB,” a final assessment described below is recommended.

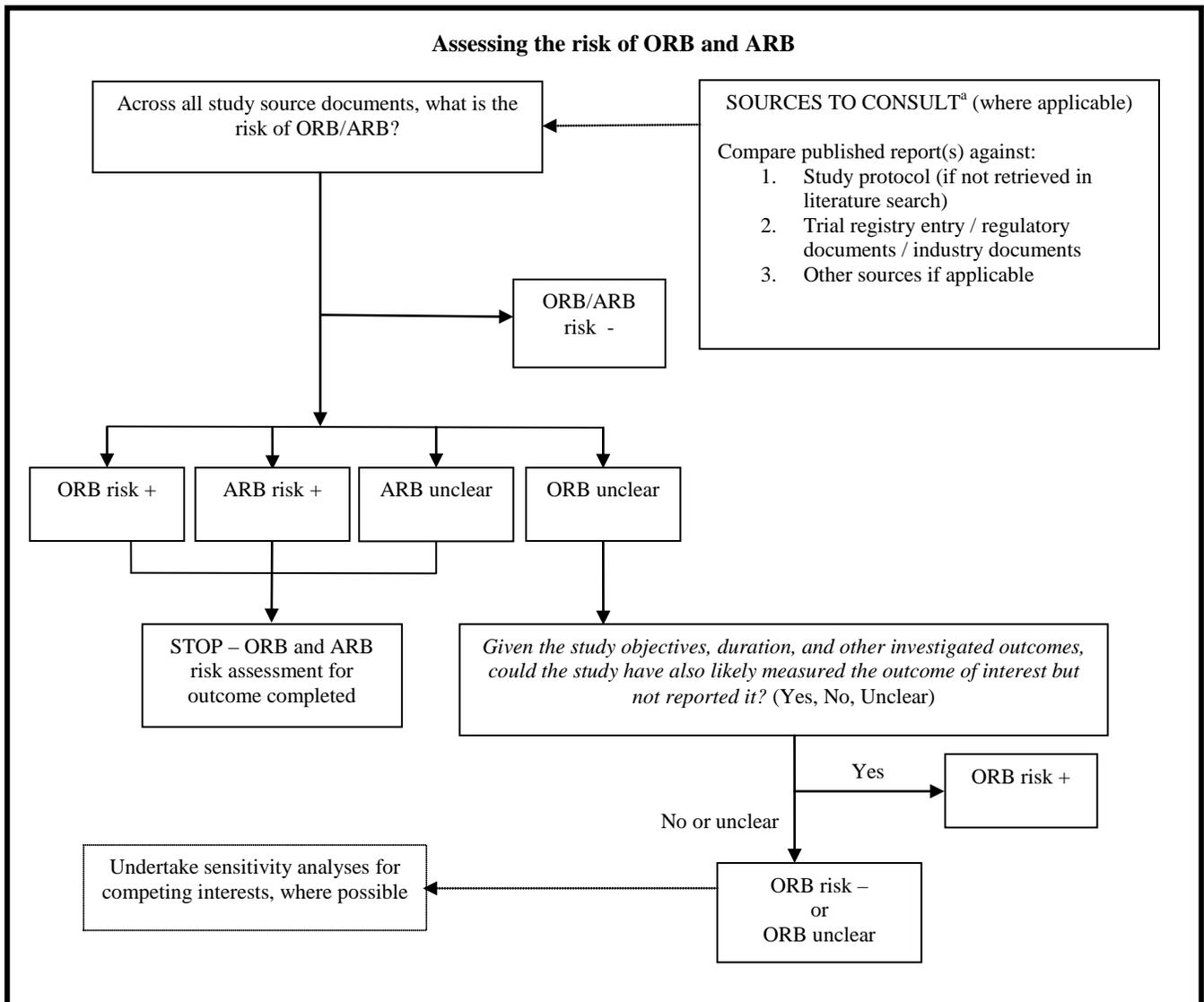
For studies for which the risk of reporting bias cannot be ruled out, we suggest that EPCs do one final assessment (Figure 1). Reviewers should ask the question: “*Given the study objectives, duration, and other investigated outcomes, could the study have also likely measured the outcome of interest but not reported it?*” If the answer is “no” the study should be rated as “ORB risk–”. 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 +” 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 Kirkham et al.²³ and illustrated by Dwan, et al.¹¹⁵ which uses a multistep process that reviewers can use to determine if potentially eligible trial reports are prone to reporting bias (available at <http://www.trialsjournal.com/content/11/1/52/table/T1>). Briefly, the matrix:

- Includes all included studies (accompanied with all corresponding publications) irrespective of whether or not 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.
- Arranges outcomes in columns and studies in rows for all included studies. The outcomes tabulated include all the review-relevant outcomes as well as outcomes that are not of review interest but are reported in included studies.

- Should differentiate complete, partial, and nonreporting for each review-relevant outcome for which the risk of reporting bias is being assessed

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.

Combining Studies 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 a substantial 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.¹¹⁶⁻¹¹⁸ Since data on only about 1,600 out of 4,600 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.

Assessing for Publication Bias

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 studies (presumably negative) 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 recommended a very cautious and judicious approach to statistical testing for the lateral asymmetry of funnel plot.¹¹⁹ Sensitivity analyses can assess whether a finding of treatment benefit is robust to differing assumptions regarding the extent of potential bias.¹²⁰⁻¹²² 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 sensitivity analyses are undertaken, reviewers should discuss how findings influence their confidence on review findings. 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,⁶ 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 of 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 Preferred Reporting Items for Systematic Reviews and Meta-Analyses 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 reporting bias. Each review requires a thoughtful, individualized approach to identifying selective outcome and analysis reporting, which must be outlined in the review, along with the rationale for that approach. Most important, 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 a completed review.¹¹⁵ 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 outcome and analysis reporting bias 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 selective outcome or analysis reporting, describe what that specific information was and how it contributed to the identification of selective outcome or analysis reporting, and the assessment of reporting bias.
- 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.
- Finally, it will often happen that systematic reviewers will find themselves with documentation about a trial from various sources, containing varying degrees of conflicting detail. Since we cannot know which source is the more accurate, we recommend that authors of systematic reviews report when such discrepancies occur and report whether the results of sensitivity analyses suggest differences in results depending on which sets of data are included.

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. Psaty BM, Kronmal RA. Reporting mortality findings in trials of rofecoxib for Alzheimer disease or cognitive impairment: a case study based on documents from rofecoxib litigation. *JAMA*. 2008 Apr 16;299(15):1813-7. PMID: 18413875.
3. 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.
4. 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.
5. Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database Syst Rev*. 2007(2):MR000005. PMID: 17443628.
6. Relevo R, Balslem H. Finding evidence for comparing medical interventions: Agency for Healthcare Research and Quality (AHRQ) and the Effective Health Care program. *J Clin Epidemiol*. 2011 Jun 16;64(11):1168-77. PMID: 21684115.
7. 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.
8. 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.
9. Sterling TD. Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa. *J Am Stat Assoc*. 1959;54(285):30-4.
10. Sterling TD, Rosenbaum WL, Weinkam JJ. Publication decisions revisited: the effect of the outcome of statistical tests on the decision to publish and vice versa. *Am Stat*. 1995:108-12.
11. Simes RJ. Publication bias: the case for an international registry of clinical trials. *J Clin Oncol*. 1986;4(10):1529-41. PMID: 3760920.
12. Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA*. 1990;263(10):1385-9. PMID: 2406472.
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. PMID: 18769481.
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. PMID: 19067477.
18. Veitch E. Tackling publication bias in clinical trial reporting. *PLoS announces the launch of a new online journal*. *PLoS Med*. 2005;2(10):e367. PMID: 17523250.
19. Sterne JAC, Egger M, Moher D. Chapter 10: Addressing reporting biases. In: Higgins JPT, 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. PMID: 15161896.
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. Turner EH, Knoepflmacher D, Shapley L. Publication bias in antipsychotic trials: An analysis of efficacy comparing the published literature to the US Food and Drug Administration database. *PLoS Med*. 2012 Mar;9(3):e1001189. PMID: 22448149.
25. 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. PMID: 18199864.
26. 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.
27. 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.
28. De Angelis C, Drazen JM, Frizelle FA, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *Ann Intern Med*. 2004;141(6):477-8. PMID: 15355883.
29. Zarin DA, Tse T, Williams RJ, et al. The ClinicalTrials.gov results database—update and key issues. *N Engl J Med*. 2011 Mar 3;364(9):852-60. PMID: 21366476.
30. Higgins JTP, Green S, eds. *Cochrane handbook for systematic reviews of interventions*. Version 5.1.0 [updated March 2011] ed: The Cochrane Collaboration; 2011.
31. 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.
32. Food and Drug Administration Modernization Act of 1997. 21 USC 301. U.S.A; 1997.
33. Food and Drug Administration Amendments Act of 2007. 21 USC 301. U.S.A; 2007.
34. Turner EH. Closing a loophole in the FDA Amendments Act. *Science*. 2008 Oct 3;322(5898):44-6. PMID: 18832629.
35. ClinicalTrials.gov. <http://www.clinicaltrials.gov>. Accessed Feb 1, 2013.
36. World Health Organization. International Clinical Trials Registry Platform (ICTRP). 2012. <http://apps.who.int/trialsearch/AdvSearch.aspx>. Accessed May 10, 2012.
37. 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.
38. Loder E, Groves T, Macauley D. Registration of observational studies. *BMJ*. 2010;340:c950. PMID: 20167643.
39. Should protocols for observational research be registered? *The Lancet*. 2010;375(9712):348. PMID: 20113809.
40. 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.
41. Sorensen HT, Rothman KJ. The prognosis for research. *BMJ*. 2010;340:c703. PMID: 20164129.
42. Wood AJ. Progress and deficiencies in the registration of clinical trials. *N Engl J Med*. 2009;360(8):824-30. PMID: 19228628.

43. 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.
44. 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.
45. Viergever RF, Ghersi D. The quality of registration of clinical trials. *PLoS One*. 2011;6(2):e14701. PMID: 21383991.
46. Shamliyan T. Reporting of results of interventional studies by the information service of the National Institutes of Health. *Clin Pharmacol*. 2010;2:169-76. PMID: 22291502.
47. 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.
48. 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.
49. Ross JS, Mulvey GK, Hines EM, et al. Trial publication after registration in ClinicalTrials.gov: a cross-sectional analysis. *PLoS Med*. 2009;6(9):e1000144. PMID: 19901971.
50. 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.
51. Golder S, Loke YK, Bland M. Unpublished data can be of value in systematic reviews of adverse effects: methodological overview. *J Clin Epidemiol*. 2010 May 8;63(10):1071-81. PMID: 20457510.
52. 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.
53. 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 Med*. 2009 January 01, 2009;6(1):e17.
54. 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.
55. 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.
56. 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.
57. 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.
58. 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.
59. 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.
60. 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.
61. 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.
62. 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.

63. O'Connor AB. The need for improved access to FDA reviews. *JAMA*. 2009 Jul 8;302(2):191-3. PMID: 19584349.
64. Doshi P, Jefferson T. The First 2 Years of the European Medicines Agency's Policy on Access to Documents: Secret No Longer. *Arch Intern Med*. 2012 Dec 19;112. PMID: 23255144.
65. European Medicines Agency. European Medicines Agency policy on access to documents (related to medicinal products for human and veterinary use). London: 2010. http://www.ema.europa.eu/docs/en_GB/document_library/Other/2010/11/WC500099473.pdf.
66. Steinbrook R. The European Medicines Agency and the Brave New World of Access to Clinical Trial Data. *Arch Intern Med*. 2012 Dec 19;112. PMID: 23254180.
67. 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
68. 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. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6_R1/Step4/E6_R1_Guideline.pdf. Accessed April 13, 2012
69. ClinicalTrials.gov Glossary of Common Site Terms. <http://clinicaltrials.gov/ct2/info/glossary>. Accessed April 13, 2012.
70. Horton R. Pardonable revisions and protocol reviews. *Lancet*. 1997 Jan 4;349(9044):6. PMID: 8988113.
71. McNamee D, James A, Kleinert S. Protocol review at The Lancet. *Lancet*. 2008 Jul 19;372(9634):189-90. PMID: 18640443.
72. Godlee F. Publishing study protocols: making them visible will improve registration, reporting and recruitment. *BMC News and Views*. 2001;2:4.
73. 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.
74. Chan A-W, Kroleza-Jeric K, Schmid I, et al. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ*. 2004;171(7):735-40. PMID: 15451835.
75. 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.
76. 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.
77. Van Royen P, Sandholzer H, Griffiths F, et al. Are presentations of abstracts at EGPRN meetings followed by publication? *Eur J Gen Pract*. 2010;16(2):100-5. PMID: 20504264.
78. 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. *J Clin Oncol*. 2008 May 1;26(13):2205-11. PMID: 18445846.
79. Chokkalingam A, Scherer R, Dickersin K. Agreement of data in abstracts compared to full publications. *Controlled Clinical Trials*. 1998;19(3):S61-S2-S2.
80. Hopewell S, Clarke M, Askie L. Reporting of trials presented in conference abstracts needs to be improved. *J Clin Epidemiol*. 2006;59(7):681-4. PMID: 16765270.
81. 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):1281-7. PMID: 16537738.
82. Shamlivan T, Kane RL. Clinical research involving children: registration, completeness, and publication. *Pediatrics*. 2012 May;129(5):e1291-300. PMID: 22529271.

83. 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.
84. Mullan RJ, Flynn DN, Carlberg B, et al. Systematic reviewers commonly contact study authors but do so with limited rigor. *J Clin Epidemiol*. 2008 Nov 13;62(2):138-42. PMID: 19013767.
85. Young T, Hopewell S. Methods for obtaining unpublished data. *Cochrane Database of Systematic Reviews*. John Wiley & Sons, Ltd; 1996.
86. Kyzas PA, Ioannidis JP. Selective reporting biases in cancer prognostic factor studies. *J Natl Cancer Inst*. 2005 Jul 20;97(14):1043-55. PMID: 16030302.
87. Butler D. Drug firm to share raw trial data. *Nature*. 2012 Oct 18;490(7420):322. PMID: 23075958.
88. Coombes R. GlaxoSmithKline grants researchers access to clinical trial data. *BMJ*. 2012;345:e6909. PMID: 23065357.
89. Clinical Trials at Novo Nordisk. <http://www.novonordisk.com/website/search/trial-result.aspx>. Accessed January 17, 2013.
90. Jefferson T, Doshi P, Thompson M, et al. Ensuring safe and effective drugs: who can do what it takes? *BMJ*. 2011;342:c7258. PMID: 21224325.
91. U.S. Cochrane Center. Resources for handsearchers | US Cochrane Center. 2011. <http://us.cochrane.org/resources-handsearchers>. Accessed April 21, 2012.
92. 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).
93. 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. *Res Synth Methods*. 2012.
94. 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.
95. Egger M, Zellweger-Zähner T, Schneider M, et al. Language bias in randomised controlled trials published in English and German. *Lancet*. 1997;350(9074):326-9. PMID: 9251637.
96. 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.
97. 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.
98. 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 Technol Assess*. 2003;7(1):1-76. PMID: 12583822.
99. 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.
100. Jüni P, Holenstein F, Sterne J, et al. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *Int J Epidemiol*. 2002;31(1):115-23. PMID: 11914306.
101. 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.
102. 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.

103. 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. *J Clin Epidemiol.* 2005;58(8):769-76.e2—76.e2. PMID: 16086467.
104. 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.
105. 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.
106. 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.
107. 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.
108. 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.
109. 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.
110. Gotzsche PC, Jorgensen AW. Opening up data at the European Medicines Agency. *BMJ.* 2011;342:d2686. PMID: 21558364.
111. Cordoba G, Schwartz L, Woloshin S, et al. Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. *BMJ.* 2010;341:c3920. PMID: 20719825.
112. 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.
113. Schierbeck LL, Rejnmark L, Tofteng CL, et al. Effect of hormone replacement therapy on cardiovascular events in recently postmenopausal women: randomised trial. *BMJ.* 2012;345:e6409. PMID: 23048011.
114. Schroll J, Lundh A. Was the composite outcome specified in the original protocol? *BMJ.* 2012 Dec 3;345:e8144. PMID: 23208256.
115. 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.
116. Eyding D, Lelgemann M, Grouven U, et al. Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials. *BMJ.* 2010;341:c4737. PMID: 20940209.
117. McGauran N, Wieseler B, Kreis J, et al. Reporting bias in medical research - a narrative review. *Trials.* 2010;11(1):37. PMID: 20388211.
118. Wieseler B, McGauran N, Kaiser T. Finding studies on reboxetine: a tale of hide and seek. *BMJ.* 2010;341:c4942. PMID: 20940211.
119. 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.
120. 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.
121. 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.
122. 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
CONSORT	CONsolidated Standards of Reporting Trials
CER	Comparative Effectiveness Review
EPC	Evidence-based Practice Center
FDA	U.S. Food and Drug Administration
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
NIH	National Institutes of Health
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

Appendix A. Definitions of the Data Elements From ClinicalTrials.gov (www.clinicaltrials.gov)

Field Name	Definition of the Data Element
NCT ID	The ClinicalTrials.gov identifier
Other IDs	Other identification numbers assigned to the protocol, including unique identifiers from other registries and NIH grant numbers
Title	Official name of the protocol provided by the study principal investigator or sponsor
Acronym	Acronym or initials used to identify this study
Funded By	Funding source as industry, NIH, U.S. Federal Government, Network, or other
Sponsors	Name of primary organization that oversees implementation of study and is responsible for data analysis
Recruitment	# Enrolling by invitation: participants are being (or will be) selected from a predetermined population # Active, not recruiting: study is ongoing (i.e., patients are being treated or examined), but participants are not currently being recruited or enrolled # Completed: the study has concluded normally; participants are no longer being examined or treated (i.e., last patient's last visit has occurred) # Suspended: recruiting or enrolling participants has halted prematurely but potentially will resume # Terminated: recruiting or enrolling participants has halted prematurely and will not resume; participants are no longer being examined or treated # Withdrawn: study halted prematurely, prior to enrollment of first participant
Conditions	Primary disease or condition being studied, or focus of the study. Diseases or conditions should use the National Library of Medicine's Medical Subject Headings (MeSH) controlled vocabulary when possible.
Study Types	Interventional or observational studies
Study Designs	Purpose, phase, treatment allocation, masking of the treatment status; type of primary outcome or endpoint that the protocol is designed to evaluate
Phases	Phase of investigation, as defined by the U.S. FDA for trials involving investigational new drugs
Study Results	<ul style="list-style-type: none"> - Participant Flow - Baseline Characteristics - Outcome Measures and Statistical Analyses - Adverse Events Information - Administrative Information
Interventions	<ul style="list-style-type: none"> - Drug (including placebo) - Device (including sham) - Biological/Vaccine - Procedure/Surgery - Radiation - Behavioral (e.g., Psychotherapy, Lifestyle Counseling) - Genetic (including gene transfer, stem cell and recombinant DNA) - Dietary Supplement (e.g., vitamins, minerals)
Outcome Measures	Specific key measurement(s) or observation(s) used to measure the effect of experimental variables in a study, or for observational studies, to describe patterns of diseases or traits or associations with exposures, risk factors or treatment.
Gender	Physical gender of individuals who may participate in the protocol
Age Groups	Age of participants
Enrollment	Number of subjects in the trial
First Received	Date the protocol information was received
Start Date	Date that enrollment to the protocol begins
Completion Date	Final date on which data was (or is expected to be) collected
Last Updated	Date the protocol information was updated
Last Verified	Date the protocol information was last verified

Field Name	Definition of the Data Element
Primary Completion Date	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
Why Study Stopped?	A brief explanation of why suspended, terminated or withdrawn studies have been halted or terminated

Appendix B. Definitions of the Data Elements From the World Health Organization International Clinical Trials Registry Platform

www.who.int/ictrp/network/trds/en/index.html

Field Name	Definition of the Data Element
Primary Registry	Name of Primary Registry
Primary Registry ID	Unique ID number assigned by the Primary Registry to this trial
Date of Registration in Primary Registry	Date when trial was officially registered in the Primary Registry.
Secondary Identifying Numbers	<ul style="list-style-type: none"> • The Universal Trial Number • Identifiers assigned by the sponsor • Other trial registration numbers issued by other Registries • Identifiers issued by funding bodies, collaborative research groups, regulatory authorities, ethics committees / institutional review boards, etc.
Source(s) of Monetary or Material Support	Major source(s) of monetary or material support for the trial (e.g. funding agency, foundation, company, institution)
Primary Sponsor	The individual, organization, group or other legal entity which takes responsibility for initiating, managing and/or financing a study.
Secondary Sponsor(s)	Additional individuals, organizations or other legal persons, if any, that have agreed with the primary sponsor to take on responsibilities of sponsorship.
Contact for Public Queries	Email address, telephone number and postal address of the contact who will respond to general queries, including information about current recruitment status.
Contact for Scientific Queries	The PI may delegate responsibility for dealing with scientific enquiries to a scientific contact for the trial. This scientific contact will be listed in addition to the PI.
Public Title	Title intended for the lay public in easily understood language.
Scientific Title	Scientific title of the study as it appears in the protocol submitted for funding and ethical review.
Countries of Recruitment	The countries from which participants will be, are intended to be, or have been recruited at the time of registration.
Health Condition(s) or Problem(s) Studied	Primary health condition(s) or problem(s) studied (e.g., depression, breast cancer, medication error).
Intervention(s)	For each arm of the trial record a brief intervention name plus an intervention description. For drugs: generic name, or temporary identifier for drugs that do not yet have a generic name; for other types of interventions: a brief descriptive name.
Key Inclusion and Exclusion Criteria	Inclusion and exclusion criteria for participant selection, including age and sex.
Study Type	Study type consists of: <ul style="list-style-type: none"> • Type of study (interventional or observational) • Study design including: <ul style="list-style-type: none"> ○ Method of allocation (randomized/non-randomized) ○ Masking (is masking used and, if so, who is masked) ○ Assignment (single arm, parallel, crossover or factorial) ○ Purpose • Phase (if applicable)
Date of First Enrollment	Anticipated or actual date of enrolment of the first participant.
Target Sample Size	Number of participants that this trial plans to enroll in total

Field Name	Definition of the Data Element
Recruitment Status	Recruitment status of this trial: <ul style="list-style-type: none"> • Pending: participants are not yet being recruited or enrolled at any site • Recruiting: participants are currently being recruited and enrolled • Suspended: there is a temporary halt in recruitment and enrolment • Complete: participants are no longer being recruited or enrolled • Other
Primary Outcome(s)	For each primary outcome the name of the outcome, the metric or method of measurement used, and the timepoint(s) of primary interest.
Key Secondary Outcomes	Secondary outcomes with the same description as primary outcomes (above).

Appendix C. FDA Web Site – Drugs@FDA

Agency	URL	Description
U.S. Food and Drug Administration: Drugs@FDA	http://www.accessdata.fda.gov/scripts/cder/drugsatfda/	<p>Drugs@FDA allows you to search for official information about FDA approved brand name and generic drugs and therapeutic biological products currently approved for sale in the United States.</p> <p>Includes the following:</p> <ul style="list-style-type: none"> • monoclonal antibodies • cytokines, growth factors, enzymes, immunomodulators; and thrombolytics • proteins intended for therapeutic use that are extracted from animals or microorganisms, including recombinant versions of these products (except clotting factors) • other nonvaccine therapeutic immunotherapies <p>Does not include:</p> <ul style="list-style-type: none"> • over-the-counter products approved for marketing through a process other than submission of a New Drug Application or Biologic License Application • drugs sold outside the United States that are not approved for marketing in the U.S. • drugs not approved by the FDA • drugs under review at FDA for which no action (approved or not approved) has occurred yet • dietary supplements • biological products regulated by the Center for Biologics Evaluation and Research • animal drugs{Center for Drug Evaluation and Research, 2010 #1164}

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

Appendix D. Grant Databases

Database	Search Terms?	Data Downloadable?	Grant Number?	Listed Publications?	Provided Registration Status?	Allowed Results Posting?	Comprehensive When Compared to Other Sources?
NIH RePORTER http://report.nih.gov/index.aspx	Yes	Yes	Yes	Yes but not accurate	No	No	Yes
Foundation Directory Online (FDO)	Yes	Yes	Yes	No	No	No	No
HSRProj http://wwwcf.nlm.nih.gov/hsr_project/home_proj.cfm	No	Yes	Yes	No	No	No	No
AHRQ GOLD http://gold.ahrq.gov/projectsearch/	No	Yes	Yes	Yes but not accurate	No	No	No