AHA SCIENTIFIC STATEMENT

Methodological Standards for Meta-Analyses and Qualitative Systematic Reviews of Cardiac Prevention and Treatment Studies

A Scientific Statement From the American Heart Association

ABSTRACT: Meta-analyses are becoming increasingly popular, especially in the fields of cardiovascular disease prevention and treatment. They are often considered to be a reliable source of evidence for making healthcare decisions. Unfortunately, problems among meta-analyses such as the misapplication and misinterpretation of statistical methods and tests are long-standing and widespread. The purposes of this statement are to review key steps in the development of a meta-analysis and to provide recommendations that will be useful for carrying out meta-analyses and for readers and journal editors, who must interpret the findings and gauge methodological quality. To make the statement practical and accessible, detailed descriptions of statistical methods have been omitted. Based on a survey of cardiovascular meta-analyses, published literature on methodology, expert consultation, and consensus among the writing group, key recommendations are provided. Recommendations reinforce several current practices, including protocol registration; comprehensive search strategies; methods for data extraction and abstraction; methods for identifying, measuring, and dealing with heterogeneity; and statistical methods for pooling results. Other practices should be discontinued, including the use of levels of evidence and evidence hierarchies to gauge the value and impact of different study designs (including meta-analyses) and the use of structured tools to assess the quality of studies to be included in a meta-analysis. We also recommend choosing a pooling model for conventional meta-analyses (fixed effect or random effects) on the basis of clinical and methodological similarities among studies to be included, rather than the results of a test for statistical heterogeneity.

Goutham Rao, MD, FAHA, Chair
Francisco Lopez-Jimenez, MD, MSc, FAHA, Vice Chair
Jack Boyd, MD
Frank D'Amico, PhD
Nefertiti H. Durant, MD, MPH
Mark A. Hlatky, MD, FAHA
George Howard, DrPH, FAHA
Katherine Kirley, MD, MS
Christopher Masi, MD, PhD
Tiffany M. Powell-Wiley, MD, MPH, FAHA
Anthony E. Solomonides, DPhil, MMath
Colin P. West, MD, PhD
Jennifer Wessel, PhD
On behalf of the American Heart Association Council on Lifestyle and Cardiometabolic Health; Council on Cardiovascular and Stroke Nursing; Council on Cardiovascular Surgery and Anesthesia; Council on Clinical Cardiology; Council on Functional Genomics and Translational Biology; and Stroke Council

Despite the increasing popularity of meta-analyses and systematic reviews in general, problems with methodology are widespread and frequently undermine the credibility of the results. New guidance is needed for both researchers who carry out meta-analyses and systematic reviews in general and the consumers who read them and rely on the results. The term meta-analysis was coined in 1976 by the American statistician Gene Glass, who wrote, "I use it to refer to the statistical analysis of a large collection of results from individual studies for the purpose of integrating findings. It connotes a rigorous alternative to the casual, narrative discussions of research studies which typify our attempts to make sense of the rapidly expanding literature."[1] Meta-analyses are a subcategory of the broader category of studies known as systematic reviews.

Qualitative systematic reviews include explicit and detailed methods for identification, selection, and grading the quality of individual studies and overall evidence but do not pool results across studies. Meta-analysis is synonymous with the term quantitative systematic review and by definition includes pooling of results across studies. The emphasis in this statement is on meta-analysis because that is the area in which methodological problems have been best documented. However, our recommendations also apply to systematic reviews more broadly.

Meta-analysis has become incredibly popular. There are now >10 000 meta-analyses and qualitative systematic reviews published annually, roughly double the number published annually just 5 years ago.[2] Meta-analyses have become especially common in the broad cardiovascular field. A PubMed search with the terms meta-analysis and cardiovascular in the title/abstract yielded 53 results for the year 2000, 413 for the year 2010, and 1196 for the year 2014.

The fundamental appeal of meta-analysis, which partly explains its popularity, is the idea of integrating evidence from multiple sources to provide reliable answers to important questions. The evidence-based medicine movement in general promotes a systematic approach to assessing the quality of evidence, considering not only research design but also other characteristics of individual articles.[3] Some evidence-based medicine hierarchies of evidence, however, assign the highest level (quality) of evidence to systematic reviews of randomized trials.[4] Meta-analyses (and systematic reviews in general) are therefore sometimes automatically accorded a great deal of credibility. The placement of meta-analyses of randomized trials (and individual randomized trials) at the top of evidence hierarchies is controversial because some believe that it underestimates the value of other research designs and because detailed assessment of methodology is not taken into account.[5,6] Indeed, questions related to the use of meta-analyses have seldom focused on methodology but rather on issues of uptake and implementation such as what strategies are most effective in encouraging more use of systematic reviews among clinicians and policy makers.[7] This is truly unfortunate because problems with published meta-analyses, including unstandardized methods and misapplication and misinterpretation of statistical and other techniques, are widespread and long-standing.[8] Bailar,[9] for example, wrote in 1997, "In my own review of select meta-analyses, problems were so frequent and so serious, including bias on the part of the meta-analyst, that it was difficult to trust the overall best estimates the method produces." More recently, Berlin and Golub[10] expressed concern that the problem of poor quality has only increased as the number of meta-analyses has grown. Searching the literature has become easier, and statistical packages for meta-analyses are freely available, so authors with limited expertise can produce poor-quality meta-analyses with ease. Flawed meta-analytic methodology is common in many fields such as oncology and from many sources, including the highly respected Cochrane Collaboration.[11,12] Furthermore, as pointed out by Ioannidis[13] in an article published in 2016, the number of meta-analyses and systematic reviews in general has grown exponentially, many are unnecessary, and there has been little improvement in recent years in their methodological quality.[13]

Despite important reports on optimal systematic review methodology and evolutions in the field,[14] tools to evaluate the quality of meta-analyses, which could guide their development, are lacking. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement was developed to address the problem of poor reporting, especially missing information, in published systematic reviews.[15] The statement includes a checklist for items that should be reported but does not address which methods should be applied and pays little attention to statistical methodology. The Assessing the Methodological Quality of Systematic Reviews (AMSTAR) checklist is a similar reporting checklist with a greater focus on methodology, including statistical methods.[16]

This American Heart Association statement serves different purposes for different audiences. First, our goal is to provide guidance to those who wish to carry out meta-analyses, especially in the fields of cardiovascular disease prevention and treatment. Second, we believe that this statement will also be useful for consumers of meta-analyses who wish to assess methodological quality, not just the completeness of reporting. Finally, we believe that the statement will be useful for journal editors who must decide whether to publish a particular meta-analysis. We believe that those who have at least some knowledge and experience in carrying out systematic reviews and meta-analyses are best

prepared to take advantage of this scientific statement. However, to make this statement broadly accessible, we have deliberately omitted mathematical formulas and technical jargon, providing a relevant reference when necessary. With increasing emphasis on patient-centered outcomes research, it can be argued that a variety of stakeholders, including patients and caregivers, as well as clinicians and methodologists, should be involved in the development of meta-analyses. Making our statement as easy to understand as possible was therefore a priority.

We describe important concepts briefly and provide a glossary of key terms, but we do not attempt to provide a comprehensive guide to performing a meta-analysis. Most meta-analyses are completed with the use of software for compiling and analyzing data from individual studies. A detailed review of available software programs is outside the scope of this scientific statement. Many programs are available free of charge. Stawicki[17] provides a brief review.

We have assumed that the reader has already formulated a sound clinical question to be addressed[18] and has substantiated the need to address the question through meta-analysis. Substantiating the need is an extremely important first step, given the huge volume of meta-analyses published annually and the questionable value of many.[3] This statement is written to inform the stepwise approach to developing a meta-analysis. We include a description of the important step of registering the protocol for a meta-analysis. We then address 9 important questions explicitly: (1) What are effective methods for searching for studies to include in a meta-analysis? (2) How should studies be selected for inclusion? (3) What are acceptable methods for data extraction and standardization from individual studies? (4) How should quality of individual studies be assessed? (5) How should heterogeneity be quantified and handled? (6) What are acceptable methods for pooling results across studies, and how do these methods vary according to study design and the frequency of outcomes? (7) What are acceptable methods for identifying publication bias? (8) What are acceptable methods and guiding principles for carrying out sensitivity and subgroup analyses? (9) Finally, what are emerging meta-analytical methods for studies addressing cardiovascular prevention and treatments?

Our key recommendations are based on a survey and review of published cardiovascular meta-analyses, a broad range of methodological literature, the consensus of the writing group, and further consultation with statisticians and others in the field. The writing group consists of a diverse group of physician–scientists, statisticians, and an expert in medical informatics experienced in developing, teaching about, and editing meta-analyses and systematic reviews in general.

SURVEY OF PUBLISHED CARDIOVASCULAR META-ANALYSES

To better inform the recommendations in this article, we carried out a survey of recently published meta-analyses in the broad cardiovascular field. Our intent was to gain insight into the methods being applied rather than gauging the quality of those methods. We searched PubMed using the terms *meta-analyses* and *cardiovascular* in the title/abstract limited to humans, English language publications, then core clinical journals (where we believed the highest-quality meta-analyses were likely to be published), and the period of January 1, 2014, to October 1, 2015. Two members of the writing group (G.R. and C.M.) carefully reviewed the methods sections of all articles independently. The citations, search methods, type of meta-analysis, inclusion/exclusion criteria for individual studies, use of quality assessment tools, pooling methods, methods for evaluating heterogeneity, use of subgroup, sensitivity analysis, and meta-regression, methods for detection of publication bias, and type of software used were extracted from each article and stored in a searchable database. Any disagreements between reviewers were resolved by discussion and consensus.

We retrieved a total of 117 citations. Thirty-five were considered irrelevant because they were not actually meta-analyses or because they did not include ≥1 cardiovascular prevention or treatment outcomes, leaving 82 relevant meta-analyses (Data Supplement). Thirteen used individual patient-level data methods only; 1 used individual patient-level data and conventional methods; 1 was purely a network meta-analysis; 2 used both bayesian and network methods; 1 was purely a bayesian meta-analysis; 1 used both conventional and network methods; 1 used both bayesian and conventional methods; 2 used both dose-response and conventional methods; 1 used bayesian, network, and conventional methods; and 59 used conventional methods only. Search methods for individual studies were fairly uniform across meta-analyses. Meta-analyses specified inclusion of experimental, observational, or both categories of studies. Use of tools to assess individual study quality was common: 18 meta-analyses used the Cochrane Risk of Bias Tool for randomized trials, and 21 used the Newcastle-Ottawa Scale for observational studies. A variety of statistical methods were applied across meta-analyses, with no consistency or any description of why specific methods (eg, pooling methods) were chosen.

PROTOCOL REGISTRATION AND SEARCH METHODS
Protocol Registration

A protocol for conducting a meta-analysis should be written and registered before the work is initiated. A

rigorously planned protocol can help investigators anticipate problems, complete reviews more efficiently, and remove bias introduced through decisions made in response to the data. A protocol can also enhance consistency among study team members when extracting and using data from primary research. Arbitrary decisions on which studies to include and how to use them are less likely to occur when guidelines are available to the review team.[19,20] Bias resulting from selective reporting is less likely to occur when methods and analytic strategies are clearly described.[21,22] In addition, protocols can serve as precedents for review strategies, thereby reducing duplication of efforts by other study teams.

The most widely used international registry for systematic review protocols is the International Prospective Register of Systematic Reviews (PROSPERO).[23] By completing 22 required fields, authors enter key protocol information, including title, review questions, types of studies to be included, data extraction (selection and coding), quality assessment, strategy for data synthesis, and analysis of subgroups or subsets. Authors are encouraged to update their protocol entries whenever changes are made during the review process.[24] By doing so, investigators contemplating a systematic review can determine whether a similar review is underway or has been completed. In addition, readers can compare published results with information in the registry to determine whether protocol deviations may have led to bias. For these reasons, journals are increasingly requiring protocol registration information as part of the submission process.[26]

Search Strategy: Sources

The search strategy should be reported in detail. A reasonable search strategy for cardiovascular or other studies begins with PubMed, a database provided by the National Library of Medicine's National Center for Biotechnology Information. Expanding the search to include Embase, Scopus, or the Cochrane Central Register of Controlled Trials can improve identification because these databases do not perfectly overlap[27] and relying solely on National Library of Medicine databases can miss 30% to 80% of relevant studies.[28,29] Depending on the question being addressed, the search can be supplemented with other sources such as trial registries and trial results registries (eg, US Food and Drug Administration, International Standard Randomized Controlled Trial Number Register, Pharma websites, or ClinicalTrials.gov), Internet searches (eg, Google Scholar), bibliographies from other studies, subject-specific databases (eg, International Pharmaceutical Abstracts), citation indexes (eg, Web of Science), data repositories and institutional in-house studies, the Turning Research into Practice database, "gray" literature sources for unpublished studies (eg, conference abstracts or trial registries), other need-specific databases (eg, Health InterNetwork Access to Research Initiative (HINARI) for low-income country access), the International Network for the Availability of Scientific Publications, and region-specific databases (eg, African Index Medicus). Hand searching involves a manual review of the entire contents of relevant journals or conference proceedings and, when feasible and focused appropriately, can sometimes increase the yield of relevant abstracts and articles compared with electronic searching alone. Similarly, manually reviewing the reference lists of articles retrieved through electronic searches for additional relevant articles is another useful way to supplement an electronic search. What is critically important rather than the specific sources searched is that, in addition to searching the essential National Library of Medicine and similar databases, the meta-analysis team should clearly specify the rationale for searching other sources or explain the reasons for not expanding a basic search to include other sources.

Search Strategy: Study Eligibility

Criteria defining study eligibility should be established before the search is initiated to enhance reproducibility and to minimize bias in the search and selection of studies. Eligibility criteria typically include study design, year of publication, languages, sample size, duration of follow-up, similarity of treatments or exposure, similarity of outcomes, and completeness of information.[30]

It is common to find several publications from the same study population. If >1 publication is based on the same cohort or population and reports the same outcomes, only the most recent or comprehensive publication should be included. In some cases, however, 2 (or more) studies may be based on the same cohort but report different outcomes or different subgroups. In such cases, it is reasonable to include both.

Some have suggested excluding very small studies. It may be more appropriate, however, to include all studies and use sample size as a grouping variable for prespecified subgroup analyses. With regard to eligibility based on similarity of treatments or outcomes, there is a tradeoff between face validity and overgeneralization, depending on the main research focus and question to be answered. For example, studies focused on a very specific therapy (eg, a study assessing the effect of metoprolol in the treatment of heart failure with decreased left ventricular ejection fraction, excluding other β-blockers) will have greater homogeneity but less generalizability, whereas a study focused on a medication class or a group of similar therapies might be more comprehensive but risks "comparing apples and oranges." Studies with incomplete information, as defined in the protocol for the meta-analysis, should generally

be excluded from analyses, assuming efforts to obtain necessary data from authors have been unsuccessful. It is recommended that such studies appear in the table as initially eligible studies with an indication of the absence of usable information. That way, the readers will understand that these studies were not missed in the search.[31]

Ideally, the decision about eligibility of studies should be made while investigators are blinded to the results of the study, the source of publication, and the institution where the study was performed. This is recommended because investigators selecting studies may be more likely to include publications that report results consistent with their preconceived ideas and may be more likely to choose studies published in prestigious journals or conducted by specific investigators.

To identify relevant studies in a particular area, it is important to create an appropriate list of key terms. These terms should include different ways to define the patient population intervention, comparators, and outcomes. For example, a meta-analysis on percutaneous coronary interventions in the elderly should include different descriptors of percutaneous coronary interventions (eg, angioplasty, coronary stenting) and different descriptors to capture studies performed in older people, under the domain elderly (eg, elderly, geriatric). If the result of the search yields only a handful of publications, the search strategy may require revision because the terms may have been too restrictive. In contrast, if the search yields thousands of publications, the search strategy may need to be modified to be more specific. Documenting key elements of each search is important to ensure reproducibility of the search and to prevent duplication of activities. Elements to document include the databases searched; years covered by the search; the date the search was run; the complete search strategy with details on search terms used; a 1- or 2- sentence summary of the search strategy, including any hand searching and review of reference lists; language restrictions; and name of the database host system.[32] These key items can be included in bullet-point form if word restrictions are an issue. Many published systematic reviews do not include this essential information.[33] Given the myriad information sources available today and the need to identify all relevant studies, we recommend enlisting the assistance of an experienced librarian with formal training in electronic literature searching, who could ideally join the study team.

SELECTION OF STUDIES AND DATA ABSTRACTION

Selection of Studies

Just as there is no specific search strategy suitable for every meta-analysis, there is no universal approach for which studies to include or exclude from a meta-analysis. Consequently, it is important for the study selection process to be explicit, consistent, and reliable. The meta-analysis protocol should include a precise description of how studies are to be selected. When there is disagreement among authors as to whether an article should be included or which data should be extracted, the first step is to revisit the study protocol to see if the disagreement is related to a lack of clarity within the protocol. As a general principle, the methods by which studies are selected for a meta-analysis, like the search methods, should be described so clearly that they are completely reproducible by an external party. It is reasonable to amend a protocol to provide more clarity. A mechanism should be established in the overall study protocol to settle persistent disagreements about study selection. For example, if 2 study team members are in disagreement about a particular study, a third team member may be called in to review the study and make a final decision. A large number of disagreements among study team members may necessitate major revisions to the study selection protocol.

In general, we recommend that review of articles proceed in 2 stages, with discussion of articles at each stage. The first stage involves review of titles and/or abstracts only; the second stage includes review of the full text of articles. Disagreements before the second stage should favor inclusion to ensure that all potentially eligible studies are assessed at the full-text review stage. This approach makes measures of agreement less useful at these earlier stages of a review, but they remain appropriate for full-text reviews. A recently published meta-analysis by Bratton et al[34] on treatment of sleep apnea and blood pressure provides an example of a precise, comprehensive description of methods for study selection.

Data Abstraction

The study protocol should also include specific guidelines for data abstraction (extraction) that include how and what type of data will be included from each study (authors, year of publication, design, statistical methods used, intervention, sample size, outcomes, etc). The protocol should describe how missing data are to be handled. For example, the study team may wish to systematically contact authors (eg, through e-mail or by telephone) of specific articles to obtain missing data. At least 2 individuals should abstract data, each working independently, and the protocol should specify how disagreements are to be resolved between them (eg, by having a third study team member make a final decision). It is desirable that authors test their data abstraction protocols through calibration exercises (ie, discussing and refining their data abstraction results among author pairs) before the formal data abstraction. Doing

so can solve major differences in data abstraction and improve their interobserver agreement before they formally perform the abstraction of data. If possible, data abstractors should be blinded to the location and any other pertinent information of each study to make the extraction as objective as possible. One of the abstractors, ideally someone with statistical expertise, should supervise and monitor the process. The type of data abstracted will depend on the purpose and type of meta-analysis but at a minimum should include basic study characteristics such as authors, year of publication, design, intervention, sample size, and outcomes. A recently published meta-analysis on low-density lipoprotein–lowering therapy and risk of prostate cancer by Tan et al[35] includes an excellent description of appropriate data abstraction:

Data from each study were independently extracted by two reviewers (Tan & Wei) using a standardized data-extraction form. Any disagreements were resolved by consensus or by consultation with a third reviewer (Yang). The following information was checked for each article: first author's last name, year of publication, location of study, study period, type of study design, mean follow-up time, drugs studied, duration of statin use, study population, number of male subjects, mean age of population, number of total cases of PCa [prostate cancer], advanced (defined by the stage of the disease as "regional" or "distant" or the TNM [tumor, node, metastasis] stage within T3-4, N1-3 and M1) and localized PCa cases (defined by the stage of the disease as 'localized' or the TNM stage as T1-2, $N_{0/x}$ and $M_{0/x}$.), high (Gleason sum ≥ 7) and low grade PCa cases (Gleason sum <7), PCa cases occurring during short- and long-term statins use ("long-term" was defined as ≥5 years of use; "short-term" was defined as <5 years of use), risk estimates [including relative risk (RR), odds ratio (OR) and hazard ratio (HR)] adjusted for the maximum number of confounding variables with corresponding 95% confidence intervals (CIs). In addition, we also tried to contact authors via e-mail to obtain further information that had not been reported in their published articles.[35]

ASSESSMENT OF STUDY QUALITY

Many meta-analyses incorporate formal methods of evaluation of individual study quality and use standardized tools such as checklists. Unfortunately, there is no uniform definition of quality for either intervention trials or observational studies. Some tools assess risk of bias, or internal validity, whereas other tools emphasize other

characteristics such as sample size, which is not usually a source of bias. An important distinction among tools is an emphasis on what a study reported in a specific article versus how that study was carried out (study conduct). Our recommendation with respect to the use of quality assessment tools is based on the following considerations, supported by leading experts in the field.

First, the sheer number of tools available to assess individual study quality is enormous, which makes it difficult to build consensus and promote a consistent approach. Sanderson et al,[36] for example, identified 86 different tools for assessing quality in observational studies alone. Second, the validity of many quality assessment tools is questionable. Very few have been scientifically validated.[37] The precise purpose of many tools is unclear because some include elements of both study reporting and study content.[38] Their development is poorly described, and not surprisingly, their content is highly variable. Finally, like their validity, the reliability of quality assessment tools has been called into question. Even the widely used Cochrane Collaboration Risk of Bias tool has shown poor interrater agreement for several of its domains.[39]

Tool Selection

Guidance for tool selection is limited. For observational studies, Sanderson et al[36] recommend that tools include a small number of key domains, be as specific as possible, be a simple checklist (rather than a score or scale), and show evidence of careful development, validity, and reliability. This advice is similar to that provided by the Agency for Healthcare Research and Quality for assessing risk of bias of individual studies comparing medical interventions.[40]

Olivo et al[41] identified 21 different scales for assessing the quality of randomized trials. The majority lacked evidence of reliability and validity. The most widely used, the Jadad scale, was an exception, having been shown to be reliable in assessments of study quality.

The Grades of Recommendation, Assessment, Development and Evaluation Working Group (GRADE) was established in 2004 and has developed an approach to the assessment of the quality of overall evidence for specific outcomes and strength of related recommendations in systematic reviews and clinical practice guidelines that has been widely adopted by a number of organizations.[42] Analysis of individual study quality is needed to inform the GRADE approach. Scores for a set of randomized trials to be included, for example, are downgraded for specific flaws such as publication bias, imprecision, and high degrees of heterogeneity among them. The GRADE approach has been shown to improve the interrater reliability of assessments of overall evidence quality compared with intuitive assessments.[43] GRADE does not provide guidance on the appropriate

selection and application of statistical methods for meta-analyses.

The writing group does not endorse any specific tool or set of tools to evaluate study quality. A careful and detailed assessment of study methodology by authors familiar with various study designs, while carefully documenting methodological strengths and weaknesses, is an appropriate and acceptable strategy. Each study would ideally be assessed by a minimum of 2 authors. In general, studies with significant methodological flaws should simply be excluded from a meta-analysis. As noted, the Jaded scale has been shown to be reliable for the assessment of individual study quality, and the GRADE approach can be useful in improving the reliability of assessments of the overall quality of evidence for specific outcomes.

HETEROGENEITY

Identifying and Measuring Heterogeneity

Broadly speaking, any type of variability among the studies included in a meta-analysis can be called heterogeneity. The 3 principal types of heterogeneity can be precisely defined: clinical heterogeneity of patients, interventions, and outcomes; methodological heterogeneity in study design (eg, double-blinded versus single-blinded studies); and statistical heterogeneity in the observed effects extracted from each study. Statistical heterogeneity is present when the observed effects extracted from each study are more variable than one would expect because of random error or chance. Normally, when the term heterogeneity appears by itself in a meta-analysis, it refers to statistical heterogeneity.[44]

Clinical heterogeneity and methodological heterogeneity are subjectively evaluated. We recommend that outcomes from individual studies not be pooled with the use of conventional meta-analytic techniques if either or both forms of heterogeneity are judged to be substantial. In such cases, proceeding with a qualitative systematic review is appropriate. For example, if 5 trials of aspirin versus placebo for the prevention of stroke vary considerably with respect to the dose of aspirin, duration of follow-up, age of patients, use of blinding, etc, it may be best to avoid pooling results for a meta-analysis. If, on the other hand, these types of differences vary only slightly among the included studies, pooling of outcomes across studies would be appropriate.

Identifying statistical heterogeneity can be informative in gauging the significance of the results of a meta-analysis. Imagine a meta-analysis of a new therapy for stroke prevention that found a stroke risk reduction of between 32% and 36% in each of 5 studies, yielding an average reduction of 34%. This result is more credible than finding an average risk reduction of 34% based on 2 studies with a 10% risk reduction and 3 studies with a 50% risk reduction. In the first case, we can have

greater confidence that the intervention is likely to have a similar effect in different settings and perhaps different populations. The observed variation among study outcomes has 2 components: the between-studies variation and the variation resulting from within-study or random error. When the between-studies variation is much greater than the within-study variation, statistical heterogeneity is said to be present.

The simplest way to identify statistical heterogeneity is through visual inspection of the forest plot of studies. If in such a graphical representation there is considerable overlap among confidence intervals (CIs) of individual study outcomes, significant statistical heterogeneity is unlikely. Statistical heterogeneity is more likely to be significant when there is little or no overlap in CIs (Figure 1). Although this informal method is useful, statistical heterogeneity should also be quantified and described more precisely. Two statistical tests are commonly used (τ^2, less commonly used, is described in Table 1).

The Cochran Q is calculated as the weighted sum of squared differences between individual study outcomes and the pooled outcome across all studies. Q is compared with the χ^2 distribution to test for significance and is therefore typically accompanied by a P value. When P is significant (<0.05), the null hypothesis that all studies share a common effect size is rejected, and significant statistical heterogeneity is said to be present. Unfortunately, Q is strongly dependent on the number of studies included in a meta-analysis and has low power when there are few studies, as is the case in most meta-analyses. Although a statistically significant Q strongly suggests that there are real differences among the study effect sizes, a nonsignificant Q should not be interpreted as a lack of real differences among the study effect sizes, particularly when the number of studies included is small.[48] Another significant disadvantage of Q is that it tells us only whether or not there is statistical heterogeneity, but not the magnitude of any statistical heterogeneity.

To more clearly quantify statistical heterogeneity, Higgins et al[49] proposed a new measure, known as inconsistency or I^2, which is the proportion of the total variation accounted for by variation in effect size between studies. Essentially, I^2 is the percentage of total variation across studies that is attributable to heterogeneity rather than chance. Despite its designation as a "square," it is possible to calculate a negative value of I^2, in which case it is set to zero. As rough guidelines, Higgins suggests threshold values of I^2 of 25%, 50%, and 75% as representing low, moderate, and high degrees of inconsistency or statistical heterogeneity, respectively. In Figure 1, for example, in the forest plot on the left, I^2 is zero, meaning statistical heterogeneity is minimal. In contrast, on the right, I^2 is 54%, indicating a moderate degree of statistical heterogeneity.
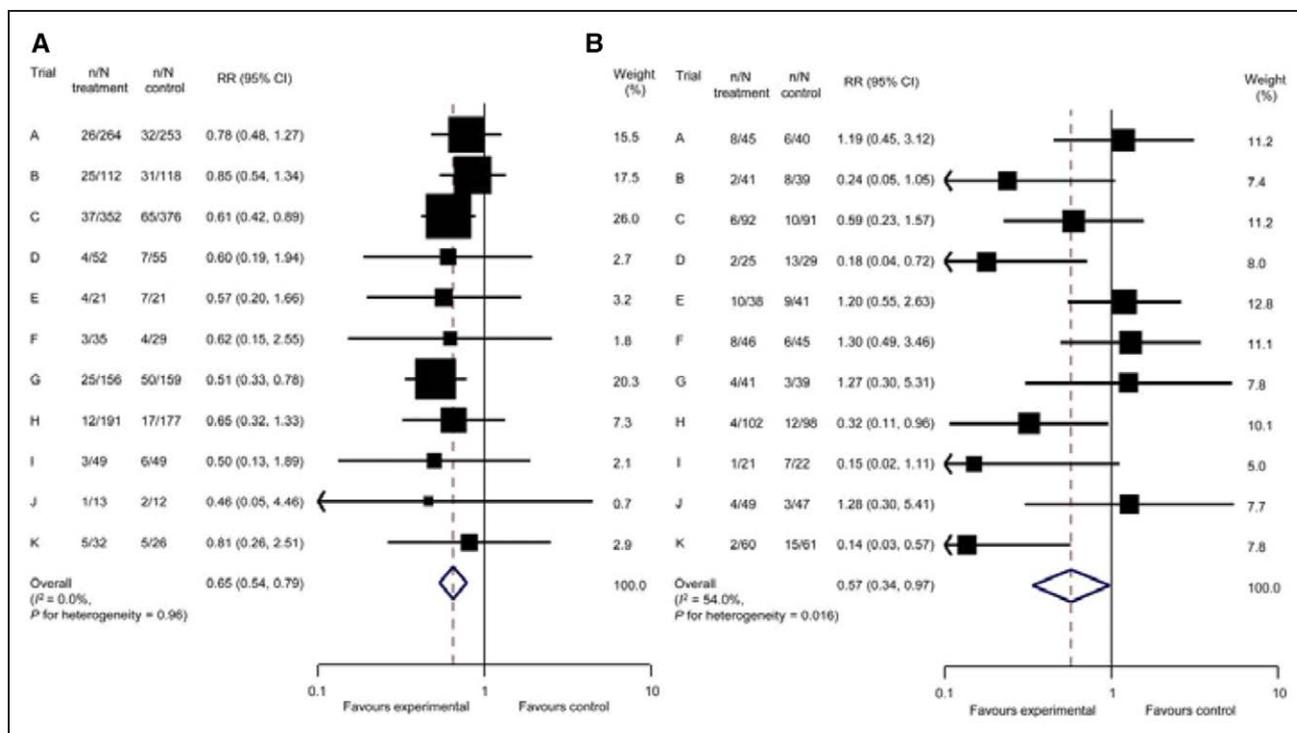
Figure 1. Forest plots revealing no significant heterogeneity and significant heterogeneity.
A, Forest plot revealing homogeneity among 11 trials of the impact of streptokinase on mortality in myocardial infarction. **B**, Forest plot revealing considerable heterogeneity among 11 studies of the impact of acetylcysteine in reducing contrast-induced nephropathy in patients undergoing angiography. CI indicates confidence interval; and RR, relative risk. Reprinted from da Costa and Juni[45] by permission of the European Society of Cardiology. Copyright © 2014, The Author.

Higgins and Thompson[50] have also proposed alternative, and much less commonly reported, measures of heterogeneity, including the H^2 statistic (I^2 is a simple transformation of H^2) and the R statistic, with the same advantages over the Q statistic.

Borenstein et al[51] recommend that an informative description of statistical heterogeneity includes both a measure of magnitude and a measure of uncertainty or significance, a recommendation that we support. Magnitude may be estimated with I^2. Uncertainty or significance may be expressed with Q and its accompanying P value or by using CIs for I^2.

Addressing Heterogeneity

After the types and extent of heterogeneity have been defined, an approach to addressing heterogeneity and its impact on a meta-analysis should be chosen. Substantial heterogeneity can suggest that quantitative pooling of results not be performed at all. The choice of a fixed-effect or random-effects pooling model (see Fixed Effect or Random Effects?) should not be based on estimated statistical heterogeneity but rather on the extent and impact of both clinical and methodological heterogeneity and whether the included studies used sufficiently similar populations, interventions, and methods.[52] If substantial clinical, methodological,

or statistical heterogeneity is present, there are 2 options: choosing not to pool data or choosing to pool data while accounting for and exploring heterogeneity.

Choosing Not to Pool Data

When clinical heterogeneity and methodological heterogeneity are significant, choosing not to pool data across studies is appropriate. Situations in which clinical heterogeneity may be too extensive to pool results include the presence of highly variable study populations such as pediatric versus adult patients or substantially different interventions. An example of methodological heterogeneity that could preclude pooling results is the mixture of randomized trials and observational studies involving substantially different populations, interventions, or procedures. In these cases, a systematic review without meta-analysis (qualitative systematic review) is the most appropriate way to summarize results.

Even if clinical heterogeneity and methodological heterogeneity do not appear to be substantial, a large degree of statistical heterogeneity across studies suggests that clinical and methodological heterogeneity should be reconsidered carefully and the plan to pool results should be re-evaluated or expanded with additional analyses.[33] In the absence of clinical and methodological heterogeneity, but with the presence of statistical heterogeneity, authors have 2 options: (1) they

Table 1. Glossary of Key Terms

Relevant Section	Definition/Explanation
Introduction	Narrative review: a summary of evidence addressing a clinical topic without specific adherence to methodology intended to reduce bias.
	Systematic review: a comprehensive summary of evidence addressing a focused clinical question using rigorous, systematic, and reproducible methods.
	Meta-analysis: a quantitative pooling, combining, or aggregation of estimates from different studies, commonly applied to results of a systematic review. This is typically applied to study-level data but may be applied to individual patient-level data when available.
Selection of studies/ data extraction	Effect size (as applied to meta-analysis): a measure of the relationship between 2 variables, commonly their difference or ratio; may be called treatment effect in a comparison of a deliberate intervention with a control.
	Standardized mean difference: When all studies report on the same scale, an effect size may be quantified as a raw difference (eg, between means); however, when heterogeneous scales have been used, the raw mean difference of each study may be divided by its SD to compute the standardized mean difference, also known as the Cohen d. This sample estimate of the true population parameter is somewhat biased, so a correction factor may be applied to compute a less biased estimate known as the Hedges g.
	Risk of bias: Bias is the result of systematic rather than random error. The risk of bias in a meta-analysis refers to the degree to which systematic error affects the reported results of both individual studies and the overall meta-analysis.
	Examples of biases in individual studies: selection bias, unintended differences between the intervention and control group participants; performance bias, treatment of intervention and control groups differs beyond the exposure of interest; attrition bias, differential loss of participants from groups; detection bias, effects in the intervention and control groups are measured differently; within-study reporting bias, differential reporting of outcomes within a given study, typically favoring reporting of statistically significant results.
	Examples of biases in meta-analysis: selection bias, greater likelihood of study inclusion according to type of study, time and duration, location and population, language and journal of publication, and even abstracting service used to identify studies; reporting bias, influences affecting results included in a meta-analysis, including a tendency for negative results to be suppressed (publication bias). Positive studies have shorter completion-to-publication times (time-lag bias), are published in higher-impact journals (location bias), are likely to be multiply published (duplication bias), and attract more citations (citation bias).
Heterogeneity	Heterogeneity: any type of variability among the studies included in the analyses. There are 3 types of heterogeneity commonly considered in meta-analysis: (1) Clinical heterogeneity: variability in subjects, interventions, or outcomes among studies included. (2) Methodological heterogeneity: variability in the design or quality of studies; for example, the use of double blinding in some clinical trials included in a meta-analysis but not in others (3) Statistical heterogeneity: the observation that the observed difference in treatment effects among studies in a meta-analysis is greater than one would expect on the basis of random error; when the term heterogeneity is used by itself in the medical literature, it usually refers to statistical heterogeneity.
	Cochran Q: weighted sum of squared differences between individual study effects and the pooled effect across studies, with the weights being those used in the pooling method. Q is distributed as a χ^2 statistic with k (number of studies) minus 1 df. Q has low power as a comprehensive test of heterogeneity, especially when the number of studies is small. Conversely, Q has too much power as a test of heterogeneity if the number of studies is large
	I^2: percentage of total variation across studies resulting from heterogeneity rather than chance calculated as follows, $I^2=100\%\times(Q-df)/Q$, where Q is Cochran's heterogeneity statistic and df is the number degrees of freedom (number of studies included minus 1). A value of 0% indicates no heterogeneity, whereas values >50% are considered to represent high levels of statistical heterogeneity. Advantages of I^2 over Q include that it does not depend on the number of studies and it offers a percentage that can be interpreted.
	τ^2: RevMan meta-analysis software presents an estimate of the between-study variance in a random-effects meta-analysis known as tau squared (τ^2 or tau^2). The square root of this number (ie, τ) is the estimated SD of underlying effects across studies. τ^2 Is infrequently reported in meta-analyses.
Methods for pooling	Fixed-effects analysis: In a fixed-effect model (analysis), we assume there is 1 true effect size that underlies all studies in the analysis and that all differences in observed effects are attributable to sampling error.
	Random-effects analysis: In a random-effects model (analysis), there is an assumption that there is a distribution of true effect sizes among the studies in the analysis (ie, effect sizes may vary from study to study.) Effect sizes may vary on the basis of many factors, including the age, education, or sex of participants or the intensity of the intervention.
	Forest plot: a visual summary of a meta-analysis in which the effect size estimate, confidence interval, and weight of each study are presented along with the pooled summary effect size estimate and confidence interval.
	Cumulative meta-analysis: Data from individual studies are organized in a chronological fashion from earliest to most recent, and data at each time point are pooled so that each pooling includes all studies published to that point.[46] A recent article by Clarke et al clearly describes the advantages of this approach.[47]
Publication bias/ sensitivity analyses/ subgroup analysis	Publication bias: Publication bias is a particular form of reporting bias. It represents the tendency to emphasize positive findings, to suppress negative findings, and to report statistics selectively to bolster positive findings.
	Funnel plot: a visual representation of individual study effect size estimates on the x axis plotted against their sample sizes or precisions on the y axis. In the absence of publication bias, the effect size estimates should be evenly distributed around the pooled effect size estimate, with greater variability for studies with smaller sample sizes or less precision. This leads to an inverted funnel shape for the plotted data. In practice, the usefulness of funnel plots is often limited by insufficient numbers of studies included in a meta-analysis. Statistical evaluations of publication bias are often similarly limited.
	Sensitivity analysis: a meta-analysis or meta-analyses conducted on a range of alternative values for key protocol assumptions or choices. This is done to evaluate consistency of results across those values.
	Subgroup analysis: analysis comparing effect estimates across groups of data within a meta-analysis. This may be done to explore heterogeneity of results across variables such as study characteristics and participant groups and to estimate effects specific to such variables.
	Meta-regression: an extension of subgroup analysis evaluating the association between continuous study-level explanatory variables and effect estimates.

(Continued)

Table 1. Continued

Relevant Section	Definition/Explanation
Emerging methods	Network meta-analysis: an analysis approach allowing concurrent comparison of multiple interventions, incorporating evidence from studies evaluating elements but not necessarily the entire network of relevant interventions
	Bayesian meta-analysis: a computationally intensive approach to meta-analysis that considers prior information informing associations of interest and the uncertainty of effect measures under analysis. Bayesian meta-analysis is particularly useful for indirect or network comparisons of evidence.
	Individual patient data meta-analysis: a method of pooling individual patient data across studies to provide summary effect size estimates. This approach dramatically improves investigators' ability to assess the impact of covariates and potential confounding factors but requires close collaborative relationships among teams of researchers.

can avoid pooling data or (2) they can pool data with a random-effects model, recognizing that this model does not eliminate heterogeneity and that the calculated results should be interpreted cautiously. They could also decide to investigate heterogeneity using subgroup analysis or meta-regression. For example, the pooled analysis may be stratified according the factors defined a priori as effect modifiers. With this, the investigators can determine whether the observed heterogeneity could be explained by such factors.

Choosing a Random-Effects Model
In cases in which clinical or methodological heterogeneity is present yet is not so substantial as to make pooling results unreasonable, it is appropriate to use a random-effects model for analysis (see Fixed Effect or Random Effects?).

Exploring Heterogeneity
Subgroup Analysis
With respect to meta-analysis, subgroups are normally groups of studies organized according to study-level variables such as patient characteristics, intervention characteristics, and duration of follow-up. Ideally, studies included in a meta-analysis are so similar that such variation is minimal or nonexistent. In many cases, however, significant differences among study-level variables exist. Analysis by subgroups is an important strategy to explore statistically significant heterogeneity based on underlying study-level differences.[53–55] It is best to plan subgroup analysis a priori, meaning that it is a part of the overall protocol or plan before any statistical analysis. This is common when certain study-level characteristics are obviously likely to influence a summary estimate. For example, a meta-analysis by Briasoulis and colleagues[56] investigated the impact of antihypertensive treatment among patients >65 years of age. Results were analyzed according to 2 subgroups: studies in which antihypertensives were compared with placebo and studies in which antihypertensives were compared with other antihypertensives. Commonly, however, subgroup analysis was carried out after statistically significant heterogeneity has been identified despite efforts to avoid combining results from studies that differ greatly with respect to patient, intervention, and other characteristics. How-

ever, there are situations when a priori hypotheses of factors with a presumed strong differential effect on outcomes would justify subgroup analyses even if no statistical heterogeneity were identified.

A meta-analysis on depression and the risk of coronary heart disease by Gan and colleagues[57] provides an example. In the primary analysis, the relative risk of coronary heart disease in prospective cohort studies among depressed patients was 1.31 (95% CI, 1.19–1.45), but the I^2 was 70%, indicating substantial statistical heterogeneity. The authors then carried out a number of subgroup analyses, categorizing studies according to mean age at baseline, study location, duration of follow-up, etc. They found, for example, that in the 4 studies in which the mean duration of follow-up was ≥15 years, the mean relative risk was just 1.09 (95% CI, 0.96–1.23) compared with a mean relative risk of 1.36 (95% CI, 1.24–1.49) among studies in which the mean follow-up duration was <15 years. This result suggested that duration of follow-up was an important source of heterogeneity among the studies. This simple example illustrates how subgroup analysis can be used to identify a source of heterogeneity. It should be noted, however, that in meta-analyses in which the duration of follow-up or other major predictors of outcome such as dose of medication vary substantially among the studies included, conventional meta-analytic techniques may not be valid. Alternative, more advanced methods such as a dose-response meta-analysis are used.[58,59] Consultation with an experienced methodologist is recommended when such alternatives are considered. Formally, subgroup analysis includes 2 parts: calculating the summary outcome measure and statistical heterogeneity in each subgroup and comparing the summary outcome measure across subgroups. The methods and calculations needed are well described by Borenstein and Higgins.[53]

Subgroup analysis is of limited value when the number of studies included in a meta-analysis is small. Indeed, in many cases, there may be too few studies to carry out a subgroup analysis. Most important, as emphasized by Song and colleagues,[55] when not defined a priori, subgroup analyses should be considered exploratory and used primarily to explore potential sources of heterogeneity. Results of such analyses should be

interpreted cautiously. If, for example, an overall estimate of an intervention effect from a meta-analysis is not statistically significant, it would be inappropriate to emphasize a conclusion that the effect was significant in a specific subgroup of studies through an exploratory subgroup analysis. Differing conclusions based on individual subgroups of studies and all studies in a meta-analysis are often related to the Simpson paradox, the phenomenon in which the association between 2 dichotomous variables is similar (negative or positive) between subgroups of studies (eg, all subgroups show a benefit of a specific drug) but different when all results are pooled together. The reason may be differences in confounding variables such as treatment arm size among the studies being pooled. A detailed description of this phenomenon can be found in an article by Rucker and Schumacher.[60]

Meta-Regression

Meta-regression is a more advanced method than subgroup analysis for exploring heterogeneity. In meta-regression, the principal meta-analytic outcome measure is the dependent variable in a linear regression model that uses a study-level variable as the independent variable. Meta-regression then allows the simultaneous exploration of multiple study-level variables (covariates) as potential sources of heterogeneity.[61] It is important to note that the covariates chosen should have been defined a priori and described in the protocol. Meta-regression has greater statistical power than subgroup analysis and can detect the presence of heterogeneity even when a test for statistical heterogeneity yields a nonsignificant result (eg, a nonsignificant Cochran Q).[62]

The covariates used in meta-regression are usually study or baseline patient characteristics. Study characteristics might include the presence or type of a comparison group, blinding, and duration of treatment or follow-up. Patient characteristics might include age, sex, and presence or severity of illness. It is important to note that the covariate used in meta-regression is always a study-level variable. Each study included in the analysis contributes a single observation that is weighted according to study precision.

Barria Perez et al[62] in a meta-analysis comparing the effects of bivalirudin and heparin on outcomes after percutaneous coronary intervention used meta-regression analyses to explore heterogeneity caused by patient demographics and the dose of heparin used in the included studies. Meta-regression analysis showed that the effect of bivalirudin on the important outcome of major bleeding was greater when unfractionated heparin doses were >60 IU/kg ($\beta=-0.012$; $P=0.030$). This meta-regression was therefore useful in identifying a source of heterogeneity among the studies.

Patient characteristics should be used cautiously in meta-regression. The data used in meta-regression are aggregated at the study level. For example, to investigate the relationship between an outcome and the age of participants, the meta-regression would analyze the average age of all participants in a study, not the ages of each individual participant. The statistical power of meta-regression to explore the association between a patient characteristic and an outcome is less than the power of a meta-analysis using individual patient data.[63] Meta-regression using a patient characteristic is prone to aggregation bias, also known as the ecological fallacy. If we consider once again the example of age as a patient characteristic of interest, the result of a meta-regression analysis may indicate that an outcome worsens as mean age in each study increases. These results should not be interpreted to indicate that the outcome worsens as age increases because the mean age within each study may not accurately reflect the ages of all participants in the study and their respective chance of experiencing the outcome. An excellent example of aggregation bias is provided in an overview by Baker et al.[61]

The decision to conduct a meta-regression analysis requires a reasonable expectation that ≥1 specific characteristics of the various included studies could be a source of heterogeneity. Meta-regression should not be carried out routinely in all meta-analyses or because results of a meta-analysis suggest that an association exists that does not have a clear potential clinical or biological explanation. Choice of the covariates to be examined via meta-regression should be grounded in a clinical understanding of the problem at hand. Covariates for meta-regression should be chosen a priori whenever possible. The number of covariates examined should be minimized to avoid multiple comparisons, which increase the chance that a false-positive result will be observed. In addition, to ensure adequate statistical power, a suggested rule of thumb is that at least 10 studies are required to test 1 covariate.[62]

A random-effects model is normally appropriate for meta-regression.[62] Results of meta-regression may be depicted graphically, with the outcome graphed on the y axis and the covariate of interest graphed on the x axis. Each data point represents a study, and the size of the point should reflect the weight assigned to the study. The regression line should be depicted and the model equation provided. An example is provided in Figure 2.

Results of meta-regression should be interpreted as hypothesis generating only. Meta-regression is an observational rather than experimental approach and is therefore subject to possible confounding. The review article by Baker et al[61] includes an excellent set of recommendations for carrying out meta-regression.
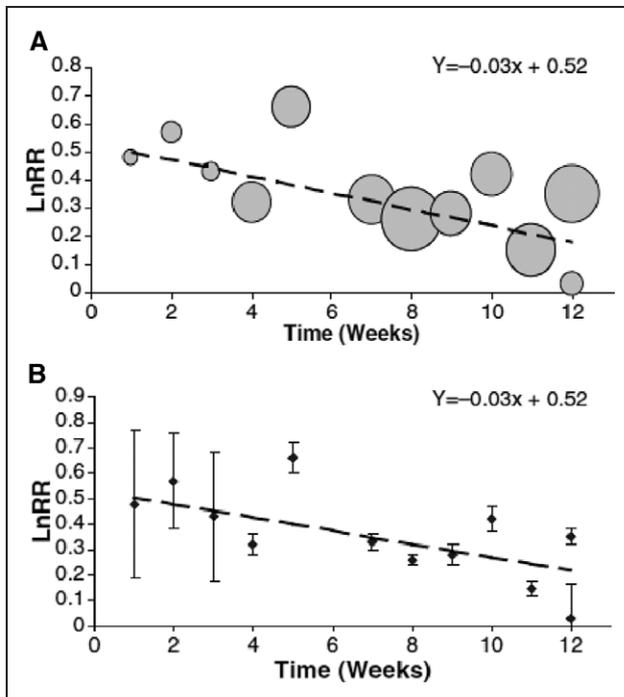
Figure 2. Graphical representation of meta-regression.
Sample meta-regression. **A**, Individual studies weighted by precision. Larger circles indicate more precise studies. **B**, The same information with point estimates of individual studies accompanied by 95% confidence intervals. Wider confidence intervals indicate less precise studies. RR indicates risk ratio. Reprinted from Baker et al[61] with permission of the publisher. Copyright © 2009, Blackwell Publishing Ltd.

POOLING METHODS

Either a random-effects or fixed-effect statistical model can be used in conventional meta-analyses. Briefly, the fixed-effect model assumes that all studies in the analysis are estimating the same common effect size. This is a reasonable assumption only when all studies included in the meta-analysis are identical or nearly identical in terms of participants and interventions (both experimental and control). The goal of the fixed-effect model is to estimate the effect size common to the studies included. In contrast, the assumption underlying the random-effects model is that the studies included are sufficiently heterogeneous that they cannot be describing the same population and that each is estimating the effect size in its respective population or setting. The random-effects model is generally more appropriate because the studies included in meta-analyses are rarely identical (ie, they are heterogeneous) with respect to patient population or intervention.

Fixed Effect or Random Effects?

As our survey of recently published cardiovascular prevention and treatment meta-analyses revealed, the

rationale for use of either a fixed-effect or random-effects model is rarely provided. In some cases, it is based on the characteristics of studies to be included in the meta-analysis and the type of estimate desired. In other cases, the choice is based on the results of tests of statistical heterogeneity. Although the choice of model is a critical decision in carrying out a meta-analysis, the guidance available to meta-analysts is limited and sometimes not definitive or clear. Higgins et al,[64] for example, have developed a tool to assess the quality of a meta-analysis. Within the tool, guidance on choosing a model is limited to the statement, "Fixed-effect meta-analysis in the presence of heterogeneity may be very misleading."

On the basis of a review of published articles, texts, and online resources, we support the guidance provided by Borenstein et al.[65] The choice of a fixed-effect model should be based on 2 important factors: whether the included studies are functionally identical, meaning they include similar or nearly identical populations, interventions, and methods, and whether the goal of synthesis of results across studies is to compute a common effect size that is applicable to populations similar or identical to those included but not generalizable to other populations. If these 2 conditions are not met, a random-effects model is appropriate.

We recommend choosing a model before carrying out a test for statistical heterogeneity. Such tests are prone to error as a result of low power. Ultimately, the most rational approach to systematic reviews and meta-analysis is to pool results across studies only if there are enough similarities and, if so, to use a fixed-effect model only if these similarities are substantial enough to support the notion that the studies are functionally identical and the goal is to estimate a single common effect size.

Guidance on Specific Statistical Pooling Methods

Beyond the choice of a fixed-effect or random-effects model, the specific pooling method and associated formula depend on the level of measurement for the dependent outcome variable. These methods are extremely well established and widely accepted by experts in meta-analysis. Our guidance is consistent with these established norms, which have been summarized by the Cochrane Collaboration[66] (Table 2).

As with all biomedical research, outcome variables in meta-analyses are usually dichotomous or continuous. Four pooling methods are commonly used for dichotomous outcomes, typically expressed as odds ratios, risk ratios, hazard ratios, or risk differences. Three of these use fixed-effect statistical models: the inverse variance method (also known as the Wolf method), the Mantel-Haenszel method, and the Peto (also known as the 1-step) method.

Table 2. Recommended Specific Pooling Methods

Type of Data	Effect Measure	Fixed-Effect Methods	Random-Effects Methods
Dichotomous	Odds ratio	Mantel-Haenszel Inverse variance Peto	Mantel-Haenszel Inverse variance
	Risk ratio	Mantel-Haenszel Inverse variance	Mantel-Haenszel Inverse variance
	Risk difference	Mantel-Haenszel Inverse variance	Mantel-Haenszel Inverse variance
Continuous	Mean difference	Inverse variance	Inverse variance
	Standardized mean difference	Inverse variance	Inverse variance
O-E and variance	User-specified (default Peto odds ratio)	Peto	None
Generic inverse variance	User specified	Inverse variance	Inverse variance
Other data	User specified	None	None

O-E indicates observed–expected. Reprinted from *Cochrane Handbook for Systematic Reviews of Interventions*[66] with permission of the publisher. Copyright © 2011, The Cochrane Collaboration.

The standard random-effects method has been the DerSimonian and Laird[67] method. The validity of this method, however, has been called into question in recent years, especially when the number of included studies is small.

Among fixed-effect methods, both the inverse variance and Mantel-Haenszel approaches are likely to yield unreliable estimates when the event under study is rare (event rates <1%). The Peto method provides reliable estimates when events are rare, intervention effects are small (odds ratios close to 1), and control and intervention groups are relatively balanced in terms of size.[68] There are circumstances, of course, when there are zero events in 1 or even both groups for specific outcomes and an odds ratio cannot be calculated. A number of methods (eg, adding a fixed value to cells with zero events) have been used to overcome this problem. Adding a fixed value to zero cells in 1 arm is likely to bias the results toward no difference between arms. More advanced methods for dealing with zero cells are described in detail by Sweeting et al.[69]

For continuous data, the standard fixed-effect method is the inverse variance method. The standard random-effects method has been the DerSimonian and Laird method. However, in recent years, for both dichotomous and continuous data, the validity of the DerSimonian and Laird method has been called into question, especially when the number of included studies is small.[70,71]

PUBLICATION BIAS

Publication bias refers to the tendency of studies with statistically significant results to be published compared with studies with nonsignificant results of meta-analysis. Publication bias is sometimes the result of editorial policies or editors' decisions. Other causes include time lag bias, that is, studies with unfavorable findings take longer to be published, and language bias, that is, articles originally written in languages other than English are more likely to report significant findings and are of lower methodological quality.[72] The significance of language bias, however, has been disputed by some.[73]

Publication bias is a common and significant issue. A formal assessment, for example, has shown strong evidence of publication bias in 10 of 28 recent meta-analyses of clinical trials and 4 of 19 recent meta-analyses of observational studies in 4 major journals (*British Medical Journal*, *Journal of the American Medical Association*, *The Lancet*, and *PLoS Medicine*).[74]

Prevention of Publication Bias

Obviously, the most effective approach to reduce the possibility of publication bias is to include all relevant studies in the meta-analysis. The US Food and Drug Administration Modernization Act of 1997 required the National Institutes of Health to create and maintain a registry of clinical trials regulated by the US Food and Drug Administration.[75] As a response, the ClinicalTrials.gov website was launched in 2000. Consistent with these efforts, the World Health Organization has recommended trial registration and launched the International Clinical Trials Registry Platform in 2007.[76] Despite some concerns about timely compliance of results reporting within trial registries,[77] registries provide a useful mechanism to identify relevant published and unpublished clinical trials and thereby reduce publication bias. Unfortunately, there is no counterpart for the meta-analysis of observational studies.

Detection of Publication Bias

In the absence of publication bias, one can assume that although larger studies should have greater precision in the estimate of the association, there should not be differences in the average magnitude of the association between the exposure and outcome for larger and smaller studies. In many cases of publication bias, small studies that show a negative or small association are less likely to be published than small studies that show a significant positive association. This principle gave rise to the strategy of plotting the magnitude of the association (most commonly the log of the odds ratio or log of the relative risk) versus the estimated precision of the study (most commonly the standard error of the estimated association) as a screening tool to identify the level of concern for publication bias. Such plots are known as funnel plots, illustrated in Figure 3.[78] In the absence of publica-
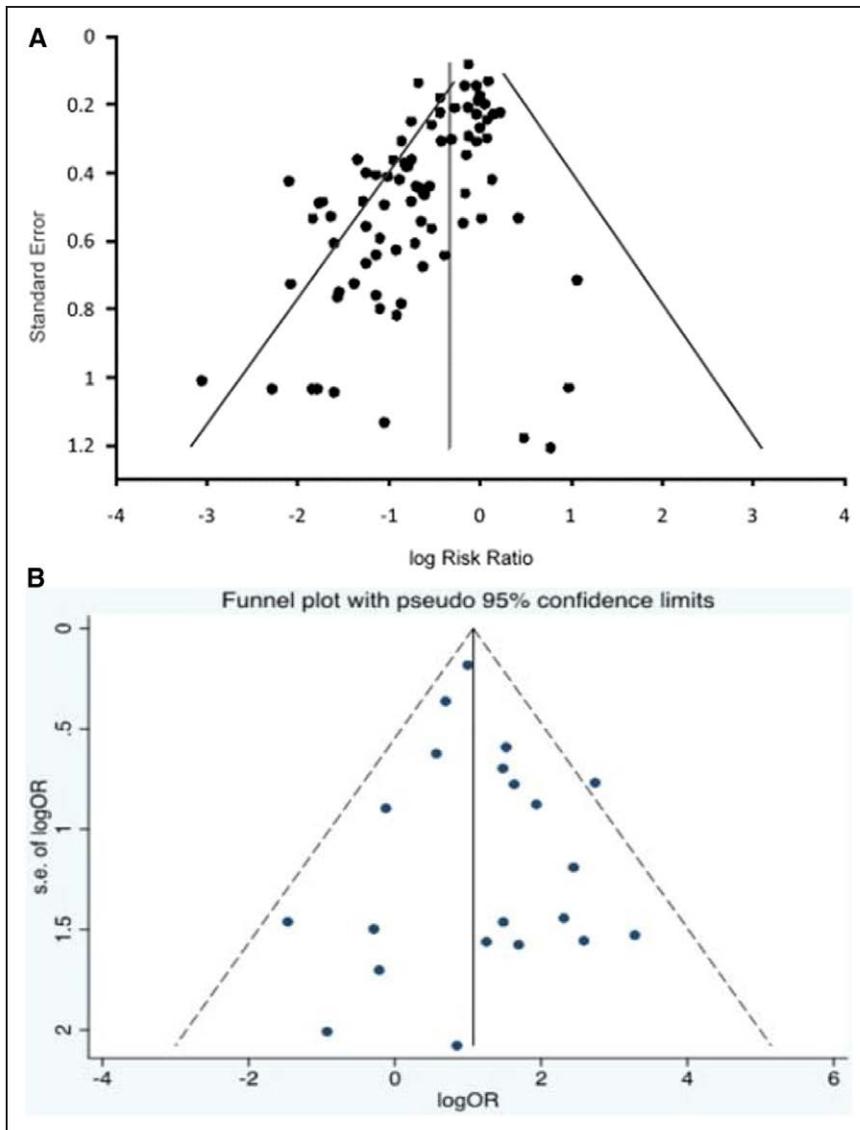
Figure 3. Asymmetric and symmetric funnel plots.
A, Asymmetric funnel plot indicating possible publication bias. Reprinted from Richie and Romanuk.[78] Copyright © 2012, The Authors. This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. **B**, Symmetric funnel plot consistent with lower likelihood of publication bias. OR indicates odds ratio. Reprinted from Liyanage et al.[79] Copyright © 2013, The Authors. This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

tion bias, the points representing the studies have a roughly symmetric funnel shape and are distributed about the average effect across the spectrum of levels of precision. In contrast, when there is publication bias, smaller, less precise studies show a significant positive effect (eg, beneficial effect of a new drug), suggesting that small negative studies were not published and leading to an asymmetric funnel. Despite their widespread use in meta-analyses, funnel plots should be interpreted carefully and have significant limitations.

A symmetric funnel plot is suggestive but does not prove the absence of publication bias, nor does an asymmetric plot prove publication bias. It is possible, for example, that the delivery of the treatment is not as well controlled in larger as in smaller studies, giving rise to a true heterogeneity of effect that leads to plots that appropriately appear asymmetric.[80] In many cases, funnel plots of subgroups of studies are symmetric but the overall funnel plot is asymmetric. In addition,

a large treatment effect can lead to a skewed sampling distribution of the log of the odds ratio or log of the relative risk, which in turn can lead to asymmetric funnel plots.[81]

Subjective interpretation of a funnel plot to detect publication bias is not ideal, so 2 statistical approaches are commonly used. The first is a rank correlation approach between the standardized estimated effect sizes (the horizontal axis of the funnel plot, standardized to the variance) and variance (the inverse of the vertical axis of the funnel plot) that was proposed by Begg and Mazumdar[82] (commonly known as the Begg test). The second approach is a regression-based method proposed by Egger and colleagues[83] in which weighted linear regression (inverse variance weighting) is used to estimate the relationship between the estimated effect sizes and the standard error of the weights (commonly known as the Egger test). The advantage of the rank correlation approach is that it does not assume a linear relationship between the effect size and its stan-

dard error, whereas the advantage of the regression approach is substantially higher power to detect asymmetry under most conditions.[81] A wealth of literature suggests modifications and improvements to both the rank correlation[84,85] and regression[86–88] approaches, with excellent reviews guiding the exact selection of the technique for the specific situation of the meta-analysis.[89,90]

Although widely used, an important limitation of funnel plots (and other methods intended to detect publication bias) deserves special attention. Specifically, the power of both subjective and analytic approaches to detect publication bias is limited when the number of studies in a meta-analysis is small. Davey and colleagues[91] described the characteristics of 22 453 meta-analyses (including 1693 meta-analyses of cardiovascular outcomes). Overall, the median number of studies included in these meta-analyses was only 3 (4 in the cardiovascular field). Ninety percent of all meta-analyses had ≤10 studies. Ninety percent of cardiovascular meta-analyses had ≤13 studies. In most meta-analyses, with either the rank correlation or regression approach, there was far less than 80% power to detect asymmetry. In >90% of the meta-analyses (and only slightly better in meta-analyses of cardiovascular diseases), the funnel plot approach and analytic approaches for detecting asymmetries were of limited value. More advanced and mathematically complex methods of dealing with publication bias (descriptions of which are beyond the scope of this article) have been described by Copas and Shi[92] and Taylor and Tibshirani.[93]

SENSITIVITY ANALYSIS

Greenhouse and Iyengar[94] define sensitivity analysis as a systematic approach to address the question of what happens if some aspect of the data or the analysis is changed. The goal of sensitivity analyses is to repeat the initial analyses by substituting alternative decisions or ranges of values for decisions that were considered arbitrary or unclear. Sensitivity analyses can be specified in advance in review protocols. However, concerns that require sensitivity analyses are often not identified until after the systematic review and meta-analysis have been completed. When sensitivity analyses show that the results of the review are not altered by different decisions made during the review or meta-analyses, the results are affirmed and considered more robust. Examples of sensitivity analyses include omitting ≥1 studies with specific patient population or intervention features or carrying out both fixed-effect and random-effects analysis and assessing the impact on summary estimates. The Cochrane Collaboration provides an excellent frame-

work in the form of guiding questions for sensitivity analysis that we recommend using as a reference[95] (Table 3).

EMERGING METHODS

Although the methods for conventional meta-analyses are well established, a number of relatively new methods are emerging. We briefly discuss 3 of these, together with their advantages and limitations. Detailed recommendations about the applications and interpretations of these methods are outside the scope of this scientific statement. We recommend, however, that before adopting these methods, study teams consult with or incorporate ≥1 individuals with expertise and experience in a specific emerging method.

Network Meta-Analyses

Network meta-analysis compares several, rather than just 2, treatments simultaneously in a single statistical model.[96] The simplest type of network meta-analysis analyzes 3 treatments (eg, A, B, and C) and calculates an indirect estimate of relative effects of 2 of the treatments (A and B) based on outcomes in trials that compare each with the same control therapy (C): A is compared with B based on data comparing A with C and B with C.[97] More complicated network models can compare ≥4 groups and incorporate both direct comparisons (ie, A to B) and indirect comparisons (A to C and B to C) in estimating the effect of treatment of A versus B.

Network meta-analyses are especially useful for evaluating cardiovascular disease treatments.[98] In contrast to pairwise comparisons, network meta-analyses have the potential to identify the best treatment for a given outcome or to provide a statistically valid estimate of the relative effects of comparable treatments. Moreover, network meta-analyses do not require lumping of similar yet distinct treatments into a single comparator of unclear utility.[99] For instance, network meta-analyses provided a method to synthesize data on several drug-eluting stent types that were studied head to head in trials of different designs and quantified differences in outcomes across stent types.[100–102] In addition, network meta-analyses can provide safety data for cardiovascular therapies in a timely manner when treatments are already used in clinical practice, as was shown in a recent network meta-analysis of smoking cessation treatments.[103]

Despite their significant strengths, especially in evaluating cardiovascular therapies, network meta-analyses are prone to important limitations. For example, the evidence from indirect and mixed comparisons may not be

Table 3. Useful Questions to Guide Sensitivity Analyses

Searching for studies
Should abstracts with results that cannot be confirmed in subsequent publications be included in the review?
Eligibility criteria
Characteristics of participants: When a majority of but not all people in a study meet an age range, should the study be included?
Characteristics of the intervention: What range of doses should be included in the meta-analysis?
Characteristics of the comparator: What criteria are required to define usual care to be used as a comparator group?
Characteristics of the outcome: What time point or range of time points are eligible for inclusion?
Study design: Should blinded and unblinded outcome assessment be included, or should study inclusion be restricted by other aspects of methodological criteria?
What data should be analyzed?
Time-to-event data: What assumptions of the distribution of censored data should be made?
Continuous data: When SDs are missing, when and how should they be imputed? Should analyses be based on change scores or on final values?
Ordinal scales: What cut point should be used to dichotomize short ordinal scales into 2 groups?
Cluster-randomized trials: What values of the intraclass correlation coefficient should be used when trial analyses have not been adjusted for clustering?
Crossover trials: What values of the within-subject correlation coefficient should be used when this is not available in primary reports?
All analyses: What assumptions should be made about missing outcomes to facilitate intention-to-treat analyses? Should adjusted or unadjusted estimates of treatment effects used?
Analysis methods
Should fixed-effect or random-effects methods be used for the analysis?
For dichotomous outcomes, should odds ratios, risk ratios, or risk differences be used?
For continuous outcomes, When several scales have assessed the same dimension, should results be analyzed as a standardized mean difference across all scales or as mean differences individually for each scale?

Reprinted from *Cochrane Handbook for Systematic Reviews of Interventions*[95] with permission of the publisher. Copyright © 2011, The Cochrane Collaboration.

valid, the rank ordering of effectiveness of treatments may be unreliable, the distribution of modifiers of the effect of an intervention may be unbalanced among studies, and the definition of best treatment may be biased or inaccurate.[104–107]

Bayesian Methods

Bayesian meta-analyses, although still a small fraction of all meta-analyses, are becoming increasingly popular. We carried out a PubMed search that revealed 267 indexed bayesian meta-analyses for 2015 compared with just 24 in 2005. Bayesian statistical methods provide a way to incorporate prior knowledge about a risk

factor or treatment into a meta-analysis. For example, knowledge of the effectiveness and safety of a drug in treating a related condition (eg, stroke) might be incorporated into a meta-analysis of using the same drug to treat a different condition (eg, myocardial infarction). This prior knowledge is incorporated into a "prior distribution" of effects that, when combined with new information from studies included in the meta-analysis, generates a "posterior distribution," which specifies a range of possible effects of the intervention or phenomenon under study.

Bayesian meta-analyses have significant strengths.[108] They allow probability statements to be made about specific outcomes such as the probability of survival with treatment A versus B. They permit incorporation of evidence from a variety of sources, including subjective clinical experience. They can easily form the foundation for decision-making frameworks and are therefore helpful in making healthcare and policy decisions. The primary criticism and limitation of bayesian meta-analysis is the subjective nature of incorporating prior beliefs and constructing prior distributions. There is not uniform agreement on how to use prior information, and different approaches can introduce biases in results and interpretation.[109]

Individual Patient-Level Data Studies

In contrast to combining summary estimates aggregated from different publications, study investigators have begun to collaborate to combine individual patient-level data and perform a pooled analysis. This approach has several distinct advantages, most notably in greatly increasing the power to examine variations in treatment outcomes according to patient characteristics.[110–112]

In a collaborative, individual patient-level data meta-analysis, the investigators from each study provide access to the individual patient-level data in their study. A defined subset of data are generally shared, comprising key baseline, treatment, and outcome data. The investigators work to harmonize the definitions and coding of important data elements such as measurements of left ventricular function, which may require recoding of data for consistency across studies. The pooled individual patient-level data can then be analyzed with powerful statistical approaches such as multivariable analysis for time-to-event outcomes with the Cox model.

Individual patient-level collaborative analysis has been used with notable success in several applications such as the Cholesterol Trialists Collaboration[113] but nevertheless faces several key barriers. Because individual patient data pooling requires cooperation among different study groups, it may be difficult to have all groups agree to participate, particularly if there

are rivalries among the investigators or among the sponsors, which may be commercial competitors. Data ownership or privacy concerns may also limit access to individual patient data, particularly if they include potentially sensitive or identifying information such as genetic markers.

The analysis of pooled individual patient data raises some unique issues in addition to concerns about the main limitations of traditional study-level meta-analyses. There are often many differences in the design and execution of the studies included, as is the case with most study-level meta-analyses, and they must be accounted for in the analysis by adjustment, stratification, or sensitivity analysis. In addition, the analysis of individual patient data is generally limited to the minimum set of variables that are common across studies, not allowing for complete adjustments.

KEY RECOMMENDATIONS

Meta-Analysis and Levels of Evidence

1. Levels of evidence based solely on study design, which often place systematic reviews in general or meta-analyses in particular at the very top of evidence pyramids, should be abandoned. Instead, a careful assessment of the methods used in a meta-analysis should be carried out to determine its risk of bias and contribution to closing important gaps in knowledge.

Rationale for this recommendation: Methodological flaws in meta-analyses are widespread. We are concerned that the arbitrary designation of systematic reviews and meta-analyses as the highest level of evidence or the top of an evidence pyramid automatically accords them a great deal of credibility and that methodological flaws may be overlooked or their impact underestimated as a result. The problems with levels of evidence and evidence hierarchies/pyramids have been well described[114] (see the introduction).

Planning

1. Planning for a meta-analysis should begin by discussing and agreeing on the need or rationale for the project among team members. This rationale should be documented in writing in the form of a protocol that includes the details of study selection, abstraction of data, models for assessing associations, and criteria for interpretation of data.

Rationale for this recommendation: This point is discussed in the introduction. Thousands of meta-analyses are being carried out and published annually. Many contribute little new information or simply do not address a question that is important or controversial.

2. For conventional meta-analyses, the initial plan must always include the possibility that pooling of data may not be appropriate and that a qualitative systematic review may be the final product.

Rationale for this recommendation: It is inappropriate to pool results from studies that are substantially clinically or methodologically heterogeneous. When data cannot be pooled because of clinical or methodological heterogeneity, the team should draw conclusions based on a qualitative assessment of the included studies (see the Addressing Heterogeneity section).

3. All meta-analytic teams should include a biostatistician or methodologist with experience in meta-analytic methods.

Rationale for this recommendation: There is a general consensus among experts that performing meta-analysis requires more than access to a software program and search tools. There are far too many poorly conducted meta-analyses in which methods have been inappropriately applied, results were incorrectly interpreted, or both. Expertise in the principles and application of meta-analytic methods is essential to produce credible results.

Search Strategies

1. Search strategies should be defined in advance, should be developed with the assistance of an experienced librarian with formal training in electronic literature searching, and should be well documented. A reasonable search strategy should at least include the National Library of Medicine databases (eg, PubMed, MEDLINE), Embase, and registries of clinical trials.

Rationale for this recommendation: We found a variety of search strategies in our review of cardiovascular meta-analyses published in core clinical journals in 2014 and 2015, which varied on the basis of the nature of the question addressed, the timeliness of currently available information, etc. Searching National Library of Medicine databases and Embase, as well as clinical trial registries, was common to almost all of these meta-analyses (see the introduction and Data Abstraction section).

Quality Assessment

1. A thorough review of appropriateness of the design and conduct of included studies is critical. Many popular tools for quality assessment or risk of bias cannot be recommended. However, the widely used Jadad scale is useful for improving the reliability of the quality assessment of individual studies. The GRADE approach is useful for improving the reliability of quality assessments of

the overall evidence to be included in systematic reviews and clinical practice guidelines.

Rationale for this recommendation: Tools to evaluate the quality of studies included in a meta-analysis are numerous, highly variable in content, often unclear in purpose, and often lacking in reliability and validity.

2. We do not recommend applying reporting guidelines such as PRISMA and Meta-Analysis of Observational Studies in Epidemiology to choose meta-analytic methodology.

Rationale for this recommendation: Reporting guidelines are designed to standardize and improve the quality of reporting, not to provide guidance on how to carry out a meta-analysis (see the introduction).

Pooling Data

1. It is inappropriate to pool data from studies that are clinically or methodologically very heterogeneous (eg, significantly different populations, differing doses of interventions, etc).

2. When pooling is considered to be reasonable on the basis of clinical and methodological homogeneity, the choice of pooling model, that is, fixed effect versus random effects, should still be based on similarities among studies to be pooled in terms of populations, interventions, exposures, and outcome measures.

Rationale for these recommendations: We make it clear in the Heterogeneity section that pooling data among clinically or methodologically heterogeneous studies is inappropriate. Combining data from very different studies (ie, mixing apples and oranges) is illogical because the summary estimate derived is not meaningful or representative of single studies that address the problem. A fixed-effect model is appropriate when there is a strong reason to believe that the studies to be pooled are all estimating the same or similar effect on the basis of the population studied, the type of intervention or exposure, etc. Many meta-analyses base the choice of model on measures of statistical heterogeneity of the outcome without considering clinical and methodological heterogeneity. This practice should be discontinued.

Statistical Heterogeneity

1. Statistical heterogeneity should be measured with ≥1 established tools (I^2, Cochran's Q [χ^2]) and reported in addition to performing visual inspection of heterogeneity with forest plots. Significant degrees of statistical heterogeneity (eg, P [Cochran's Q] <0.05, or I^2>50%) should be explored with ≥1 of subgroup analysis, sensitivity analysis, or meta-regression. In some cases, even when at the outset studies appear

to be clinically and methodologically homogeneous, exploration of statistical heterogeneity may encourage the study team to reconsider the decision to pool results across studies or to perform additional analyses to explore the sources of heterogeneity according to potential effect modifiers defined a priori.

Rationale for this recommendation: Statistical heterogeneity is a useful measure of the variability of the outcomes of individual studies. Investigating why there is statistical heterogeneity (eg, are there specific clinical and methodological aspects of different studies that may explain it?) provides consumers of meta-analysis with a more nuanced and meaningful interpretation of summary measures (see the Heterogeneity section).

Publication Bias

1. The preferred strategy to reduce or eliminate publication bias is to carry out a systematic search for all published and unpublished articles.

Rationale for this recommendation: Funnel plots and associated statistical tests for publication bias may be included in a meta-analysis but are seldom useful because of the usually small number of studies included. Methods to detect publication bias are underpowered, so although a significant result suggests the presence of publication bias, a nonsignificant result does not prove that publication bias is absent (see the Publication Bias section).

Sensitivity Analysis

1. Sensitivity analysis is a useful way to evaluate the robustness of the results of a meta-analysis to assumptions made in the process and should be carried out whenever possible with the use of the guiding questions from the Cochrane Collaboration.[75]

Rationale for this recommendation: A change in how studies are selected or data are combined, for example, can have a major impact on the results and conclusions of a meta-analysis. Sensitivity analysis, when carried out appropriately, can inform a more meaningful and credible interpretation of the results of a meta-analysis.

Emerging Methods

1. Emerging methods such as network meta-analysis and bayesian methods should be undertaken only with expert guidance.

Rationale for this recommendation: These methods are still under development. Appropriate application of emerging methods requires considerable experience and expertise.

ACKNOWLEDGMENTS

The authors acknowledge the valuable help and thoughtful reviews provided by Emanuela Taioli, PhD, and Satish Iyengar, PhD, in the preparation of this manuscript.

FOOTNOTES

Circulation is available at http://circ.ahajournals.org.

DISCLOSURES

Writing Group Disclosures

Writing Group Member	Employment	Research Grant	Other Research Support	Speakers' Bureau/ Honoraria	Expert Witness	Ownership Interest	Consultant/ Advisory Board	Other
Goutham Rao	Case Western Reserve University and University Hospitals of Cleveland	None	None	None	None	None	None	None
Francisco Lopez-Jimenez	Mayo Clinic	None	None	None	None	None	None	None
Jack Boyd	Stanford University	None	None	None	None	None	None	None
Frank D'Amico	University of Pittsburgh Medical Center–St. Margaret Hospital and Duquesne University	None	None	None	None	None	None	None
Nefertiti H. Durant	University of Alabama Birmingham School of Medicine	None	None	None	None	None	None	None
Mark A. Hlatky	Stanford University School of Medicine	None	None	None	None	None	None	None
George Howard	University of Alabama at Birmingham School of Public Health	Bayer Pharmaceutical†	None	None	None	None	None	None
Katherine Kirley	American Medical Association	None	None	None	None	None	None	None
Christopher Masi	NorthShore University HealthSystem	Pfizer, Inc*	None	None	None	None	None	None
Tiffany M. Powell-Wiley	National Institutes of Health Cardiovascular and Pulmonary Branch	NIH†	None	None	None	None	None	NIH†

(Continued)

Writing Group Disclosures Continued

Writing Group Member	Employment	Research Grant	Other Research Support	Speakers' Bureau/ Honoraria	Expert Witness	Ownership Interest	Consultant/ Advisory Board	Other
Anthony E. Solomonides	NorthShore University HealthSystem Research Institute	None	None	None	None	None	None	None
Jennifer Wessel	Indiana University	AHA Scientist Development Grant*	None	None	None	Eli Lilly & Co*	None	None
Colin P. West	Mayo Clinic	None	None	None	None	None	None	None

This table represents the relationships of writing group members that may be perceived as actual or reasonably perceived conflicts of interest as reported on the Disclosure Questionnaire, which all members of the writing group are required to complete and submit. A relationship is considered to be "significant" if (a) the person receives $10 000 or more during any 12-month period, or 5% or more of the person's gross income; or (b) the person owns 5% or more of the voting stock or share of the entity, or owns $10 000 or more of the fair market value of the entity. A relationship is considered to be "modest" if it is less than "significant" under the preceding definition.

*Modest.

†Significant.

Reviewer Disclosures

Reviewer	Employment	Research Grant	Other Research Support	Speakers' Bureau/ Honoraria	Expert Witness	Ownership Interest	Consultant/ Advisory Board	Other
Ana C. Alba	Toronto General Hospital–University Health Network (Canada)	None	None	None	None	None	None	None
Evangelos Kontopantelis	University of Manchester (UK)	None	None	None	None	None	None	None
Malcolm R. Macleod	University of Edinburgh (UK)	None	None	None	None	None	None	None
Shelley McLeod	University of Toronto (Canada)	None	None	None	None	None	None	None
Gowri Raman	Tufts Medical Center	AHRQ (received contract to conduct systematic review comparing management strategies in atherosclerotic renal artery stenosis)*; Cubist/Merck Pharmaceuticals (PI on a grant to conduct systematic review of burden of illness of MRSA in complicated skin and structure infections)*	None	None	None	None	None	None

This table represents the relationships of reviewers that may be perceived as actual or reasonably perceived conflicts of interest as reported on the Disclosure Questionnaire, which all reviewers are required to complete and submit. A relationship is considered to be "significant" if (a) the person receives $10 000 or more during any 12-month period, or 5% or more of the person's gross income; or (b) the person owns 5% or more of the voting stock or share of the entity, or owns $10 000 or more of the fair market value of the entity. A relationship is considered to be "modest" if it is less than "significant" under the preceding definition.

*Significant.

REFERENCES

1. Glass GV. Primary, secondary and meta-analysis of research. *Educ Res*. 1976;5:3–8.
2. Tebala GD. What is the future of biomedical research? *Med Hypotheses*. 2015;85:488–490. doi: 10.1016/j.mehy.2015.07.003.
3. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312: 71–72.
4. Oxford Centre for Evidence-Based Medicine. Levels of evidence. March 2009. http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/. Accessed October 2, 2015.
5. Kristiansen IS. Evidence-based medicine: time for asking critical questions? ISPOR 9th Annual European Congress. http://www.ispor.org/news/articles/june07/ebmt.asp. Accessed March 25, 2016.
6. Concato J. Study design and "evidence" in patient-oriented research. *Am J Respir Crit Care Med*. 2013;187:1167–1172. doi: 10.1164/rccm.201303-0521OE.
7. Bero LA, Jadad AR. How consumers and policymakers can use systematic reviews for decision making. *Ann Intern Med*. 1997;127:37–42.
8. Bailar JC 3rd. The practice of meta-analysis. *J Clin Epidemiol*. 1995;48: 149–157.
9. Bailar JC 3rd. The promise and problems of meta-analysis. *N Engl J Med*. 1997;337:559–561. doi: 10.1056/NEJM199708213370810.

10. Berlin JA, Golub RM. Meta-analysis as evidence: building a better pyramid. *JAMA*. 2014;312:603–605. doi: 10.1001/jama.2014.8167.

11. Lundh A, Knijnenburg SL, Jørgensen AW, van Dalen EC, Kremer LC. Quality of systematic reviews in pediatric oncology: a systematic review. *Cancer Treat Rev*. 2009;35:645–652. doi: 10.1016/j.ctrv.2009.08.010.

12. Schroll JB, Moustgaard R, Gøtzsche PC. Dealing with substantial heterogeneity in Cochrane reviews. Cross-sectional study. *BMC Med Res Methodol*. 2011;11:22. doi: 10.1186/1471-2288-11-22.

13. Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q*. 2016;94:485–514. doi: 10.1111/1468-0009.12210.

14. Eden J, Levit L, Berg A, Mortin M, eds. *Finding What Works in Health Care: Standards for Systematic Reviews*. Washington, DC: Institute of Medicine; 2011.

15. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6:e1000097. doi: 10.1371/journal.pmed.1000097.

16. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, Henry DA, Boers M. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol*. 2009;62:1013–1020. doi: 10.1016/j.jclinepi.2008.10.009.

17. Stawicki SPA. Research guide: guide to statistical software: meta-analysis applications. *Opus 12 Scientist*. 2011;5:4–5.

18. Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Med Inform Decis Mak*. 2007;7:16. doi: 10.1186/1472-6947-7-16.

19. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol*. 2009;62:e1–e34. doi: 10.1016/j.jclinepi.2009.06.006.

20. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA; PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015;4:1. doi: 10.1186/2046-4053-4-1.

21. Page MJ, McKenzie JE, Kirkham J, Dwan K, Kramer S, Green S, Forbes A. Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions. *Cochrane Database Syst Rev*. 2014:MR000035.

22. Saini P, Loke YK, Gamble C, Altman DG, Williamson PR, Kirkham JJ. Selective reporting bias of harm outcomes within studies: findings from a cohort of systematic reviews. *BMJ*. 2014;349:g6501.

23. National Institute for Health Research. PROSPERO: International Prospective Register of Systematic Reviews. http://www.crd.york.ac.uk/prospero/. Accessed July 10, 2017.

24. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, Stewart L. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Syst Rev*. 2012;1:2. doi: 10.1186/2046-4053-1-2.

25. Deleted in proof.

26. PLoS Medicine Editors. Best practice in systematic reviews: the importance of protocols and registration. *PLoS Med*. 2011;8:e1001009. doi: 10.1371/journal.pmed.1001009.

27. Suarez-Almazor ME, Belseck E, Homik J, Dorgan M, Ramos-Remus C. Identifying clinical trials in the medical literature with electronic databases: MEDLINE alone is not enough. *Control Clin Trials*. 2000;21:476–487.

28. Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ*. 1994;309:1286–1291.

29. McManus RJ, Wilson S, Delaney BC, Fitzmaurice DA, Hyde CJ, Tobias RS, Jowett S, Hobbs FD. Review of the usefulness of contacting other experts when conducting a literature search for systematic reviews. *BMJ*. 1998;317:1562–1563.

30. Petitti DB. *Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis: Methods for Quantitative Synthesis in Medicine*. New York, NY: Oxford University Press; 1994.

31. McCrae N, Purssell E. Eligibility criteria in systematic reviews published in prominent medical journals: a methodological review. *J Eval Clin Pract*. 2015;21:1052–1058. doi: 10.1111/jep.12448.

32. Higgins JPT, Green S, Cochrane Collaboration. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester, UK: Wiley-Blackwell; 2008.

33. Yoshii A, Plaut DA, McGraw KA, Anderson MJ, Wellik KE. Analysis of the reporting of search strategies in Cochrane systematic reviews. *J Med Libr Assoc*. 2009;97:21–29. doi: 10.3163/1536-5050.97.1.004.

34. Bratton DJ, Gaisl T, Wons AM, Kohler M. CPAP vs mandibular advancement devices and blood pressure in patients with obstructive sleep apnea: a systematic review and meta-analysis. *JAMA*. 2015;314:2280–2293. doi: 10.1001/jama.2015.16303.

35. Tan P, Wei S, Tang Z, Gao L, Zhang C, Nie P, Yang L, Wei Q. LDL-lowering therapy and the risk of prostate cancer: a meta-analysis of 6 randomized controlled trials and 36 observational studies. *Sci Rep*. 2016;6:24521. doi: 10.1038/srep24521.

36. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol*. 2007;36:666–676. doi: 10.1093/ije/dym018.

37. Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Control Clin Trials*. 1995;16:62–73.

38. Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*. 1999;282:1054–1060.

39. Hartling L, Ospina M, Liang Y, Dryden DM, Hooton N, Krebs Seida J, Klassen TP. Risk of bias versus quality assessment of randomised controlled trials: cross sectional study. *BMJ*. 2009;339:b4012.

40. Viswanathan M, Ansari MT, Berkman ND, Chang S, Hartling L, McPheeters LM, Santaguida PL, Shamliyan T, Singh K, Tsertsvadze A, Treadwell JR. Assessing the risk of bias of individual studies in systematic reviews of health care interventions: Agency for Healthcare Research and Quality Methods Guide for Comparative Effectiveness Reviews. March 2012. AHRQ Publication No. 12-EHC047-EF. http://www.effectivehealthcare.ahrq.gov/. Accessed June 8, 2016.

41. Olivo SA, Macedo LG, Gadotti IC, Fuentes J, Stanton T, Magee DJ. Scales to assess the quality of randomized controlled trials: a systematic review. *Phys Ther*. 2008;88:156–175. doi: 10.2522/ptj.20070147.

42. Cochrane Collaboration. *Cochrane Handbook for Systematic Reviews of Interventions*. http://handbook.cochrane.org/chapter_12/12_2_1_the_grade_approach.htm. Accessed January 17, 2017.

43. Mustafa RA, Santesso N, Brozek J, Akl EA, Walter SD, Norman G, Kulasegaram M, Christensen R, Guyatt GH, Falck-Ytter Y, Chang S, Murad MH, Vist GE, Lasserson T, Gartlehner G, Shukla V, Sun X, Whittington C, Post PN, Lang E, Thaler K, Kunnamo I, Alenius H, Meerpohl JJ, Alba AC, Nevis IF, Gentles S, Ethier MC, Carrasco-Labra A, Khatib R, Nesrallah G, Kroft J, Selk A, Brignardello-Petersen R, Schünemann HJ. The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. *J Clin Epidemiol*. 2013;66:736–742; quiz 742.e1–742e-5. doi: 10.1016/j.jclinepi.2013.02.004.

44. Deeks JD, Higgins JPT, Altman D. Chapter 9: analyzing data and undertaking meta-analysis. In: Higgins JPT, Green S, eds. *Cochrane Handbook of Systematic Reviews*. 2011. http://handbook.cochrane.org/. Accessed February 20, 2016.

45. da Costa BR, Juni P. Systematic reviews and meta-analyses of randomized trials: principles and pitfalls. *Eur Heart J*. 2014;35:3336–3345.

46. Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med*. 1992;327:248–254. doi: 10.1056/NEJM199207233270406.

47. Clarke M, Brice A, Chalmers I. Accumulating research: a systematic account of how cumulative meta-analyses would have provided knowledge, improved health, reduced harm and saved resources. *PLoS One*. 2014;9:e102670. doi: 10.1371/journal.pone.0102670.

48. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Part 4: heterogeneity. In: *Introduction to Meta-Analysis*. Chichester, UK: John Wiley and Sons; 2009:105–125.

49. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327:557–560. doi: 10.1136/bmj.327.7414.557.

50. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21:1539–1558. doi: 10.1002/sim.1186.

51. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Identifying and quantifying heterogeneity. *Introduction to Meta-Analysis*. Chichester, UK: Wiley & Sons; 2009:107–125.

52. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Fixed-effect versus random effects models. In: *Introduction to Meta-Analysis*. Chichester, UK: John Wiley and Sons; 2009:84.

53. Borenstein M, Higgins JP. Meta-analysis and subgroups. *Prev Sci*. 2013;14:134–143. doi: 10.1007/s11121-013-0377-7.

54. Sedgwick P. Meta-analyses: heterogeneity and subgroup analysis. *BMJ* 2013;346:f4040 doi: 10.1136/bmj.f4040.

55. Song F, Sheldon TA, Sutton AJ, Abrams KR, Jones DR. Methods for exploring heterogeneity in meta-analysis. *Eval Health Prof*. 2001;24:126–151. doi: 10.1177/016327870102400203.

56. Briasoulis A, Agarwal V, Tousoulis D, Stefanadis C. Effects of antihypertensive treatment in patients over 65 years of age: a meta-analysis of randomised controlled studies. *Heart*. 2014;100:317–323. doi: 10.1136/heartjnl-2013-304111.

57. Gan Y, Gong Y, Tong X, Sun H, Cong Y, Dong X, Wang Y, Xu X, Yin X, Deng J, Li L, Cao S, Lu Z. Depression and the risk of coronary heart disease: a meta-analysis of prospective cohort studies. *BMC Psychiatry*. 2014;14:371. doi: 10.1186/s12888-014-0371-z.

58. Kunz LM, Normand SL, Sedrakyan A. Meta-analysis of rate ratios with differential follow-up by treatment arm: inferring comparative effectiveness of medical devices. *Stat Med*. 2015;34:2913–2925. doi: 10.1002/sim.6530.

59. Crippa A, Orsini N. Dose-response meta-analysis of differences in means. *BMC Med Res Methodol* 2016;16:91. doi: 10.1186/s12874-016-0189-0.

60. Rücker G, Schumacher M. Simpson's paradox visualized: the example of the rosiglitazone meta-analysis. *BMC Med Res Methodol*. 2008;8:34. doi: 10.1186/1471-2288-8-34.

61. Baker WL, White CM, Cappelleri JC, Kluger J, Coleman CI; Health Outcomes, Policy, and Economics (HOPE) Collaborative Group. Understanding heterogeneity in meta-analysis: the role of meta-regression. *Int J Clin Pract*. 2009;63:1426–1434. doi: 10.1111/j.1742-1241.2009.02168.x.

62. Barria Perez AE, Rao SV, Jolly SJ, Pancholy SB, Plourde G, Rimac G, Poirier Y, Costerousse O, Bertrand OF. Meta-analysis of effects of bivalirudin versus heparin on myocardial ischemic and bleeding outcomes after percutaneous coronary intervention. *Am J Cardiol*. 2016;117:1256–1266. doi: 10.1016/j.amjcard.2016.01.015.

63. Lambert PC, Sutton AJ, Abrams KR, Jones DR. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *J Clin Epidemiol*. 2002;55:86–94.

64. Higgins JP, Lane PW, Anagnostelis B, Anzures-Cabrera J, Baker NF, Cappelleri JC, Haughie S, Hollis S, Lewis SC, Moneuse P, Whitehead A. A tool to assess the quality of a meta-analysis. *Res Synth Methods*. 2013;4:351–366. doi: 10.1002/jrsm.1092.

65. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods*. 2010;1:97–111. doi: 10.1002/jrsm.12.

66. Higgins JPT, Green S, eds. Section 9.4, summarizing effects across studies. *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration. March 2011. http://www.cochrane-handbook.org. Accessed June 8, 2016.

67. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7:177–188.

68. Greenland S, Salvan A. Bias in the one-step method for pooling study results. *Stat Med*. 1990;9:247–252.

69. Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data [published correction appears in *Stat Med*. 2006;25:2700]. *Stat Med*. 2004;23:1351–1375. doi: 10.1002/sim.1761.

70. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Stat Med*. 2002;21:589–624.

71. Guolo A, Varin C. Random-effects meta-analysis: the number of studies matters. *Stat Methods Med Res*. 2017;26:1500–1518. doi: 10.1177/0962280215583568.

72. Jüni P, Holenstein F, Sterne J, Bartlett C, Egger M. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *Int J Epidemiol*. 2002;31:115–123.

73. Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, Mierzwinski-Urban M, Clifford T, Hutton B, Rabb D. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *Int J Technol Assess Health Care*. 2012;28:138–144. doi: 10.1017/S0266462312000086.

74. Kicinski M. Publication bias in recent meta-analyses [published correction appears in *PLoS One*. 2014;9]. *PLoS One*. 2013;8:e81823. doi: 10.1371/journal.pone.0081823.

75. US National Institutes of Health. History, policies, and laws. https://clinicaltrials.gov/ct2/about-site/history. Accessed June 5, 2016.

76. Ghersi D, Pang T. En route to international clinical trial transparency. *Lancet*. 2008;372:1531–1532. doi: 10.1016/S0140-6736(08)61635-9.

77. Anderson ML, Chiswell K, Peterson ED, Tasneem A, Topping J, Califf RM. Compliance with results reporting at clinicaltrials.gov. *N Engl J Med*. 2015;372:1031–1039. doi: 10.1056/NEJMsa1409364.

78. Richie ML, Romanuk TN. A meta-analysis of probiotic efficacy for gastrointestinal diseases. *PLoS One*. 2012;7:e34938. doi: 10.1371/journal.pone.0034938.

79. Liyanage SS, Rahman B, Ridda I, Newall AT, Tabrizi SN, Garland SM, Segelov E, Seale H, Crowe PJ, Moa A, Macintyre CR. The aetiological role of human papillomavirus in oesophageal squamous cell carcinoma: a meta-analysis. *PLoS One*. 2013;8:e69238. doi: 10.1371/journal.pone.0069238.

80. Stuck, AE, Rubenstein LZ, Wieland D, Bias in meta-analysis detected by a simple, graphical test: asymmetry detected in funnel plot was probably due to true heterogeneity. *BMJ*. 1998;316:469; author reply 470–471.

81. Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J Clin Epidemiol*. 2000;53:1119–1129.

82. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics*. 1994;50:1088–1101.

83. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315:629–634.

84. Rücker G, Schwarzer G, Carpenter J. Arcsine test for publication bias in meta-analyses with binary outcomes. *Stat Med*. 2008;27:746–763. doi: 10.1002/sim.2971.

85. Schwarzer G, Antes G, Schumacher M. A test for publication bias in meta-analysis with sparse binary data. *Stat Med*. 2007;26:721–733. doi: 10.1002/sim.2588.

86. Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Stat Med*. 2006;25:3443–3457. doi: 10.1002/sim.2380.

87. Macaskill P, Walter SD, Irwig L. A comparison of methods to detect publication bias in meta-analysis. *Stat Med*. 2001;20:641–654. doi: 10.1002/sim.698.

88. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat Med*. 1999;18:2693–2708.

89. Jin ZC, Zhou XH, He J. Statistical methods for dealing with publication bias in meta-analysis. *Stat Med*. 2015;34:343–360. doi: 10.1002/sim.6342.

90. Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. *J Clin Epidemiol*. 2000;53:207–216.

91. Davey J, Turner RM, Clarke MJ, Higgins JP. Characteristics of meta-analyses and their component studies in the Cochrane Database of Systematic Reviews: a cross-sectional, descriptive analysis. *BMC Med Res Methodol*. 2011;11:160. doi: 10.1186/1471-2288-11-160.

92. Copas JB, Shi JQ. A sensitivity analysis for publication bias in systematic reviews. *Stat Methods Med Res*. 2001;10:251–265. doi: 10.1177/096228020101000402.

93. Taylor J, Tibshirani RJ. Statistical learning and selective inference. *Proc Natl Acad Sci U S A*. 2015;112:7629–7634. doi: 10.1073/pnas.1507583112.

94. Greenhouse JB, Iyengar S. Sensitivity analysis and diagnostics. In: Cooper H, Hedges LV, Valentine JC, eds. *The Handbook of Research Synthesis and Meta-Analysis*. New York, NY: Russel Sage Foundation; 2009:418.

95. Higgins JPT, Green S, eds. Section 9.7, sensitivity analysis. *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration. March 2011. http://www.cochrane-handbook.org. Accessed June 6, 2016.

96. Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. *Ann Intern Med*. 2013;159:130–137. doi: 10.7326/0003-4819-159-2-201307160-00008.

97. Caldwell DM. An overview of conducting systematic reviews with network meta-analysis. *Syst Rev*. 2014;3:109. doi: 10.1186/2046-4053-3-109.

98. Biondi-Zoccai G, Abbate A, Benedetto U, Palmerini T, D'Ascenzo F, Frati G. Network meta-analysis for evidence synthesis: what is it and why is it posed to dominate cardiovascular decision making? *Int J Cardiol*. 2015;182:309–314. doi: 10.1016/j.ijcard.2015.01.023.

99. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ*. 2005;331:897–900. doi: 10.1136/bmj.331.7521.897.

100. Palmerini T, Biondi-Zoccai G, Della Riva D, Stettler C, Sangiorgi D, D'Ascenzo F, Kimura T, Briguori C, Sabatè M, Kim HS, De Waha A, Kedhi E, Smits PC, Kaiser C, Sardella G, Marullo A, Kirtane AJ, Leon MB, Stone GW. Stent thrombosis with drug-eluting and bare-metal stents: evidence from a comprehensive network meta-analysis. *Lancet*. 2012;379:1393–1402. doi: 10.1016/S0140-6736(12)60324-9.

101. Palmerini T, Biondi-Zoccai G, Della Riva D, Mariani A, Sabaté M, Valgimigli M, Frati G, Kedhi E, Smits PC, Kaiser C, Genereux P, Galatius S, Kirtane

AJ, Stone GW. Clinical outcomes with drug-eluting and bare-metal stents in patients with ST-segment elevation myocardial infarction: evidence from a comprehensive network meta-analysis. *J Am Coll Cardiol*. 2013;62:496–504. doi: 10.1016/j.jacc.2013.05.022.

102. Palmerini T, Biondi-Zoccai G, Della Riva D, Mariani A, Sabaté M, Smits PC, Kaiser C, D'Ascenzo F, Frati G, Mancone M, Genereux P, Stone GW. Clinical outcomes with bioabsorbable polymer- versus durable polymer-based drug-eluting and bare-metal stents: evidence from a comprehensive network meta-analysis. *J Am Coll Cardiol*. 2014;63:299–307. doi: 10.1016/j.jacc.2013.09.061.

103. Mills EJ, Thorlund K, Eapen S, Wu P, Prochaska JJ. Cardiovascular events associated with smoking cessation pharmacotherapies: a network meta-analysis. *Circulation*. 2014;129:28–41. doi: 10.1161/CIRCULATIONAHA.113.003961.

104. Bafeta A, Trinquart L, Seror R, Ravaud P. Reporting of results from network meta-analyses: methodological systematic review. *BMJ*. 2014; 348:g1741.

105. Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions [published correction appears in *Int J Epidemiol*. 2013;42:919]. *Int J Epidemiol*. 2013;42:332–345. doi: 10.1093/ije/dys222.

106. Kibret T, Richer D, Beyene J. Bias in identification of the best treatment in a bayesian network meta-analysis for binary outcome: a simulation study. *Clin Epidemiol*. 2014;6:451–460. doi: 10.2147/CLEP.S69660.

107. Trinquart L, Attiche N, Bafeta A, Porcher R, Ravaud P. Uncertainty in treatment rankings: reanalysis of network meta-analyses of randomized trials. *Ann Intern Med*. 2016;164:666–673. doi: 10.7326/M15-2521.

108. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. *Stat Methods Med Res*. 2001;10:277–303. doi: 10.1177/096228020101000404.

109. Higgins JP, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Stat Soc Ser A Stat Soc*. 2009;172:137–159. doi: 10.1111/j.1467-985X.2008.00552.x.

110. Stewart LA, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data: Cochrane Working Group. *Stat Med*. 1995;14:2057–2079.

111. Simmonds MC, Higgins JP, Stewart LA, Tierney JF, Clarke MJ, Thompson SG. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. *Clin Trials*. 2005;2:209–217. doi: 10.1191/1740774505cn087oa.

112. Thompson S, Kaptoge S, White I, Wood A, Perry P, Danesh J; Emerging Risk Factors Collaboration. Statistical methods for the time-to-event analysis of individual participant data from multiple epidemiological studies. *Int J Epidemiol*. 2010;39:1345–1359. doi: 10.1093/ije/dyq063.

113. Cholesterol Treatment Trialists' (CTT) Collaboration, Fulcher J, O'Connell R, Voysey M, Emberson J, Blackwell L, Mihaylova B, Simes J, Collins R, Kirby A. Colhoun H, Braunwald E, LaRosa J, Pedersen TR, Tonkin A, Davis B, Sleight P, Franzosi MG, Baigent C, Keech A. Efficacy and safety of LDL-lowering therapy among men and women: meta-analysis of individual data from 174,000 participants in 27 randomised trials. *Lancet*. 2015;385:1397–1405.

114. Cohen AM, Stavri PZ, Hersh WR. A categorization and analysis of the criticisms of evidence-based medicine. *Int J Med Inform*. 2004;73:35–43. doi: 10.1016/j.ijmedinf.2003.11.002.

Circulation

Methodological Standards for Meta-Analyses and Qualitative Systematic Reviews of Cardiac Prevention and Treatment Studies: A Scientific Statement From the American Heart Association

Goutham Rao, Francisco Lopez-Jimenez, Jack Boyd, Frank D'Amico, Nefertiti H. Durant, Mark A. Hlatky, George Howard, Katherine Kirley, Christopher Masi, Tiffany M. Powell-Wiley, Anthony E. Solomonides, Colin P. West and Jennifer Wessel

On behalf of the American Heart Association Council on Lifestyle and Cardiometabolic Health; Council on Cardiovascular and Stroke Nursing; Council on Cardiovascular Surgery and Anesthesia; Council on Clinical Cardiology; Council on Functional Genomics and Translational Biology; and Stroke Council

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/136/10/e172

Data Supplement (unedited) at:
http://circ.ahajournals.org/content/suppl/2017/08/07/CIR.0000000000000523.DC1

Data Supplement: Relevant Cardiovascular Meta-Analyses Retrieved through Electronic Literature Search							
Study Number	Title	First author	Journal reference	PMID	Type of Meta-Analysis	*Quality/Risk of Bias Tools Used for Individual Studies	Pooling Models for Conventional Meta-Analyses
1	The impact of smoking on clinical efficacy and pharmacodynamic effects of clopidogrel: a systematic review and meta-analysis.	Zhao ZG	Heart. 2014 Feb;100(3):192-9. doi: 10.1136/heartjnl-2013-304138. Epub 2013 Jun 8. Review.	23749792	Conventional	Newcastle-Ottawa Scale for non-randomized studies.	Random effects
2	Paradoxically lower prevalence of peripheral arterial disease in South Asians: a systematic review and meta-analysis.	Sebastianski M	Heart. 2014 Jan;100(2):100-5. doi: 10.1136/heartjnl-2013-303605. Epub 2013 Jun 11. Review.	:23756656	Conventional		Random effects
3	Preoperative smoking status and postoperative complications: a systematic review and meta-analysis.	Gronkjaer M	Ann Surg. 2014 Jan;259(1):52-71. doi: 10.1097/SLA.0b013e3182911913. Review.	:23799418	Conventional		Fixed effect; Random effects
4	Effects of antihypertensive treatment in patients over 65 years of age: a meta-analysis of randomised controlled studies.	Briasoulis A	Heart. 2014 Feb;100(4):317-23. doi: 10.1136/heartjnl-2013-304111. Epub 2013 Jun 27. Review.	:23813846	Conventional	Cochrane Risk of Bias Tool	Fixed effect; Random effects
5	Orthostatic hypotension and mortality risk: a meta-analysis of cohort studies.	Xin W	Heart. 2014 Mar;100(5):406-13. doi: 10.1136/heartjnl-2013-304121. Epub 2013 Jul 22. Review.	:23878177	Conventional	Newcastle-Ottawa scale	Fixed effect; Random effects
6	The prognostic value of pre-operative and post-operative B-type natriuretic peptides in patients undergoing noncardiac surgery: B-type natriuretic peptide and N-terminal fragment of pro-B-type natriuretic peptide: a systematic review and individual patien	Rodseth RN	J Am Coll Cardiol. 2014 Jan 21;63(2):170-80. doi: 10.1016/j.jacc.2013.08.1630. Epub 2013 Sep 26. Review.	:24076282	Individual patient-level data		
7	Cholesteryl ester transfer protein polymorphisms, statin use, and their impact on cholesterol levels and cardiovascular events.	Leusink M	Clin Pharmacol Ther. 2014 Mar;95(3):314-20. doi: 10.1038/clpt.2013.194. Epub 2013 Sep 30.	:24080640	Conventional		
8	Long-term coffee consumption and risk of cardiovascular disease: a systematic review and a dose-response meta-analysis of prospective cohort studies.	Ding M	Circulation. 2014 Feb 11;129(6):643-59. doi: 10.1161/CIRCULATIONAHA.113.005925. Epub 2013 Nov 7. Review.	:24201300|PMCID:PMC3945962	Dose-response;Conventional	Newcastle-Ottawa scale	Fixed effect; Random effects
9	Cardiovascular events associated with smoking cessation pharmacotherapies: a network meta-analysis.	Mills EJ	Circulation. 2014 Jan 7;129(1):28-41. doi: 10.1161/CIRCULATIONAHA.113.003961. Epub 2013 Dec 9.	:24323793|PMCID:PMC4258065	Bayesian;Network;Conventional		Random effects

10	Meta-analysis of same versus different stent for drug-eluting stent restenosis.	Vyas A	Am J Cardiol. 2014 Feb 15;113(4):601-6. doi: 10.1016/j.amjcard.2013.10.041. Epub 2013 Nov 23.	:24342760	Conventional	Downs & Black Risk Assessment Tool	Random effects
11	Reappraisal of thienopyridine pretreatment in patients with non-ST elevation acute coronary syndrome: a systematic review and meta-analysis.	Bellemain-Appaix A	BMJ. 2014 Oct 24;349:g6269. doi: 10.1136/bmj.g6269. Review.	:25954988 |PMCID:PMC4208629	Conventional	Newcastle-Ottawa scale	Fixed effect; Random effects
12	Association of all-cause and cardiovascular mortality with prehypertension: a meta-analysis.	Huang Y	Am Heart J. 2014 Feb;167(2):160-168.e1. doi: 10.1016/j.ahj.2013.10.023. Epub 2013 Nov 6. Review.	:24439976	Conventional	US Preventive Services Task Force Guidelines and a modified checklist	Fixed effect; Random effects
13	Meta-analysis of the quantitative relation between pulse pressure and mean arterial pressure and cardiovascular risk in patients with diabetes mellitus.	Kodama S	Am J Cardiol. 2014 Mar 15;113(6):1058-65. doi: 10.1016/j.amjcard.2013.12.005. Epub 2013 Dec 25. Review.	:24462067	Conventional	Newcastle-Ottawa Scale	Random effects
14	Diabetes as a risk factor for stroke in women compared with men: a systematic review and meta-analysis of 64 cohorts, including 775,385 individuals and 12,539 strokes.	Peters SA	Lancet. 2014 Jun 7;383(9933):1973-80. doi: 10.1016/S0140-6736(14)60040-4. Epub 2014 Mar 7. Review.	:24613026	Individual patient-level data; Conventional		Random effects
15	Prehypertension and the risk of stroke: a meta-analysis.	Huang Y	Neurology. 2014 Apr 1;82(13):1153-61. doi: 10.1212/WNL.0000000000000268. Epub 2014 Mar 12.	:24623843	Conventional	Guidelines for quality of studies developed by the US Preventive Services Task Force and a modified checklist	Fixed effect; Random effects
16	Effect of angiotensin-converting enzyme inhibitors and angiotensin II receptor blockers on all-cause mortality, cardiovascular deaths, and cardiovascular events in patients with diabetes mellitus: a meta-analysis.	Cheng J	JAMA Intern Med. 2014 May;174(5):773-85. doi: 10.1001/jamainternmed.2014.348.	:24687000	Conventional	Jadad Scale	Fixed effect; Random effects
17	Vitamin D and risk of cause specific death: systematic review and meta-analysis of observational cohort and randomised intervention studies.	Chowdhury R	BMJ. 2014 Apr 1;348:g1903. doi: 10.1136/bmj.g1903. Review.	:24690623 |PMCID:PMC3972416	Conventional	Newcastle-Ottawa scale	Fixed effect; Random effects

#	Title	Author	Citation	PMID	Type	Scale	Model
18	Effect of dietary pulse intake on established therapeutic lipid targets for cardiovascular risk reduction: a systematic review and meta-analysis of randomized controlled trials.	Ha V	CMAJ. 2014 May 13;186(8):E252-62. doi: 10.1503/cmaj.131727. Epub 2014 Apr 7. Review.	:24710915 |PMCID:PMC4016088	Conventional		Random effects
19	Palm oil and blood lipid-related markers of cardiovascular disease: a systematic review and meta-analysis of dietary intervention trials.	Fattore E	Am J Clin Nutr. 2014 Jun;99(6):1331-50. doi: 10.3945/ajcn.113.081190. Epub 2014 Apr 9. Review.	:24717342	Conventional		Fixed effect; Random effects
20	Association of dietary, circulating, and supplement fatty acids with coronary risk: a systematic review and meta-analysis.	Chowdhury R	Ann Intern Med. 2014 Mar 18;160(6):398-406. doi: 10.7326/M13-1788. Review. Erratum in: Ann Intern Med. 2014 May 6;160(9):658.	:24723079	Conventional	NewCastle-Ottawa scale	Random effects
21	Minimally invasive direct coronary artery bypass graft surgery or percutaneous coronary intervention for proximal left anterior descending artery stenosis: a meta-analysis.	Deo SV	Ann Thorac Surg. 2014 Jun;97(6):2056-65. doi: 10.1016/j.athoracsur.2014.01.086. Epub 2014 Apr 26.	:24775802	Conventional		Random effects
22	Meta-analysis of time-related benefits of statin therapy in patients with acute coronary syndrome undergoing percutaneous coronary intervention.	Navarese EP	Am J Cardiol. 2014 May 15;113(10):1753-64. doi: 10.1016/j.amjcard.2014.02.034. Epub 2014 Mar 2. Review.	:24792742	Conventional		Fixed effect; Random effects
23	Fish consumption and acute coronary syndrome: a meta-analysis.	Leung Yinko SS	Am J Med. 2014 Sep;127(9):848-57.e2. doi: 10.1016/j.amjmed.2014.04.016. Epub 2014 May 4.	:24802020	Conventional	Newcastle-Ottawa scale	Random effects
24	Nut consumption in relation to cardiovascular disease risk and type 2 diabetes: a systematic review and meta-analysis of prospective studies.	Zhou D	Am J Clin Nutr. 2014 Jul;100(1):270-7. doi: 10.3945/ajcn.113.079152. Epub 2014 May 7. Review.	:24808491	Conventional	Newcastle-Ottawa scale	Fixed effect; Random effects
25	The prevalence, incidence, progression, and risks of aortic valve sclerosis: a systematic review and meta-analysis.	Coffey S	J Am Coll Cardiol. 2014 Jul 1;63(25 Pt A):2852-61. doi: 10.1016/j.jacc.2014.04.018. Epub 2014 May 7. Review.	:24814496	Conventional		Random effects
26	Nut consumption and risk of type 2 diabetes, cardiovascular disease, and all-cause mortality: a systematic review and meta-analysis.	Luo C	Am J Clin Nutr. 2014 Jul;100(1):256-69. doi: 10.3945/ajcn.113.076109. Epub 2014 May 21. Review.	:24847854	Conventional		Random effects
27	Prognostic implications of global LV dysfunction: a systematic review and meta-analysis of global longitudinal strain and ejection fraction.	Kalam K	Heart. 2014 Nov;100(21):1673-80. doi: 10.1136/heartjnl-2014-305538. Epub 2014 May 23. Review.	:24860005	Conventional		Random effects
28	Higher potency statins and the risk of new diabetes: multicentre, observational study of administrative databases.	Dormuth CR	BMJ. 2014 May 29;348:g3244. doi: 10.1136/bmj.g3244.	:24874977 |PMCID:PMC4038449	Conventional		Fixed effect; Random effects

29	Consumption of nuts and legumes and risk of incident ischemic heart disease, stroke, and diabetes: a systematic review and meta-analysis.	Afshin A	Am J Clin Nutr. 2014 Jul;100(1):278-88. doi: 10.3945/ajcn.113.076901. Epub 2014 Jun 4. Review.	:24898241 |PMCID:PMC4144102	Dose response; Conventional	Other tool	Fixed effect; Random effects
30	Clinical outcomes with Î²-blockers for myocardial infarction: a meta-analysis of randomized trials.	Bangalore S	Am J Med. 2014 Oct;127(10):939-53. doi: 10.1016/j.amjmed.2014.05.032. Epub 2014 Jun 11.	:24927909	Conventional		
31	Vitamin D and mortality: meta-analysis of individual participant data from a large consortium of cohort studies from Europe and the United States.	Schottker B	BMJ. 2014 Jun 17;348:g3656. doi: 10.1136/bmj.g3656.	:24938302 |PMCID:PMC4061380	Individual patient-level data		Random effects
32	Off-pump versus on-pump coronary revascularization: meta-analysis of mid- and long-term outcomes.	Chaudhry UA	Ann Thorac Surg. 2014 Aug;98(2):563-72. doi: 10.1016/j.athoracsur.2014.05.003. Epub 2014 Jun 24.	:24968764	Conventional	Jadad composite scale Newcastle-Ottawa Scale	Random effects
33	Electrocardiogram-based predictors of clinical outcomes: a meta-analysis of the prognostic value of ventricular repolarization.	Al-Zaiti SS	Heart Lung. 2014 Nov-Dec;43(6):516-26. doi: 10.1016/j.hrtlng.2014.05.004. Epub 2014 Jun 29.	:24988910	Conventional	Quality Assessment of Diagnostic Accuracy Studies Tool	Random effects
34	Leucocyte telomere length and risk of cardiovascular disease: systematic review and meta-analysis.	Haycock PC	BMJ. 2014 Jul 8;349:g4227. doi: 10.1136/bmj.g4227. Review.	:25006006 |PMCID:PMC4086028	Conventional	Newcastle-Ottawa scale	Random effects
35	Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data.	Holmes MV	BMJ. 2014 Jul 10;349:g4164. doi: 10.1136/bmj.g4164.	:25011450 |PMCID:PMC4091648	Individual patient-level data		Fixed effect; Random effects
36	Effect on cardiovascular risk of high density lipoprotein targeted drug treatments niacin, fibrates, and CETP inhibitors: meta-analysis of randomised controlled trials including 117,411 patients.	Keene D	BMJ. 2014 Jul 18;349:g4379. doi: 10.1136/bmj.g4379. Review.	:25038074 |PMCID:PMC4103514	Conventional	Cochrane Risk of Bias Tool	Random effects
37	Relationship of body mass index with total mortality, cardiovascular mortality, and myocardial infarction after coronary revascularization: evidence from a meta-analysis.	Sharma A	Mayo Clin Proc. 2014 Aug;89(8):1080-100. doi: 10.1016/j.mayocp.2014.04.020. Epub 2014 Jul 16.	:25039038	Conventional		Random effects

38	Cardiovascular events and bleeding risk associated with intravitreal antivascular endothelial growth factor monoclonal antibodies: systematic review and meta-analysis.	Thulliez M	JAMA Ophthalmol. 2014 Nov;132(11):1317-26. doi: 10.1001/jamaophthalmol.2014.2333. Review.	:25058694	Conventional	Cochrane Risk of Bias Tool	Fixed effect
39	Cardiovascular disease and vitamin D supplementation: trial analysis, systematic review, and meta-analysis.	Ford JA	Am J Clin Nutr. 2014 Sep;100(3):746-55. doi: 10.3945/ajcn.113.082602. Epub 2014 Jul 23.	:25057156	Bayesian; Conventional	Cochrane Risk of Bias Tool	Fixed effect; Random effects
40	Fruit and vegetable consumption and mortality from all causes, cardiovascular disease, and cancer: systematic review and dose-response meta-analysis of prospective cohort studies.	Wang X	BMJ. 2014 Jul 29;349:g4490. doi: 10.1136/bmj.g4490. Review. Erratum in: BMJ. 2014;349:5472.	:25073782 |PMCID:P MC4115152	Conventional	Newcastle-Ottawa scale	Random effects
41	Very low levels of atherogenic lipoproteins and the risk for cardiovascular events: a meta-analysis of statin trials.	Boekholdt SM	J Am Coll Cardiol. 2014 Aug 5;64(5):485-94. doi: 10.1016/j.jacc.2014.02.615. Review.	:25082583 |PMCID:P MC4443441	Individual patient-level data	Delphi Score	
42	Perioperative beta blockade in noncardiac surgery: a systematic review for the 2014 ACC/AHA guideline on perioperative cardiovascular evaluation and management of patients undergoing noncardiac surgery: a report of the American College of Cardiology/Ameri	Wijeysundera DN	Circulation. 2014 Dec 9;130(24):2246-64. doi: 10.1161/CIR.0000000000000104. Epub 2014 Aug 1. Review.	:25085964	Conventional	Cochrane Risk of Bias Tool Newcastle-Ottawa Scale	
43	Association between alcohol consumption and risk of cardiovascular disease and all-cause mortality in patients with hypertension: a meta-analysis of prospective cohort studies.	Huang C	Mayo Clin Proc. 2014 Sep;89(9):1201-10. doi: 10.1016/j.mayocp.2014.05.014. Epub 2014 Aug 1.	:25091872	Conventional	Other	Fixed effect; Random effects
44	Prognostic value of cardiac troponin in patients with chronic kidney disease without suspected acute coronary syndrome: a systematic review and meta-analysis.	Michos ED	Ann Intern Med. 2014 Oct 7;161(7):491-501. doi: 10.7326/M14-0743. Review.	:25111499	Conventional	Downs and Black Checklist	Other
45	Blood pressure-lowering treatment based on cardiovascular risk: a meta-analysis of individual patient data.	Blood Pressure Lowering Treatment Trialists' Colla	Lancet. 2014 Aug 16;384(9943):591-8. doi: 10.1016/S0140-6736(14)61212-5.	:25131978	Individual patient-level data		Fixed effect; Random effects
46	Dietary linoleic acid and risk of coronary heart disease: a systematic review and meta-analysis of prospective cohort studies.	Farvid MS	Circulation. 2014 Oct 28;130(18):1568-78. doi: 10.1161/CIRCULATIONAHA.114.010236. Epub 2014 Aug 26. Review.	:25161045 |PMCID:P MC4334131	Conventional		Fixed effect; Random effects
47	Transfusion triggers for guiding RBC transfusion for cardiovascular surgery: a systematic review and meta-analysis*.	Curley GF	Crit Care Med. 2014 Dec;42(12):2611-24. doi:	:25167086	Conventional	Cochrane Risk of Bias Tool	Random effects

			10.1097/CCM.0000000000000548. Review.			
48	Efficacy of Î² blockers in patients with heart failure plus atrial fibrillation: an individual-patient data meta-analysis.	Kotecha D	Lancet. 2014 Dec 20;384(9961):2235-43. doi: 10.1016/S0140-6736(14)61373-8. Epub 2014 Sep 2.	:25193873	Individual patient-level data	Cochrane Risk of Bias Tool	Fixed effect
49	The prognostic importance of weight loss in coronary artery disease: a systematic review and meta-analysis.	Pack QR	Mayo Clin Proc. 2014 Oct;89(10):1368-77. doi: 10.1016/j.mayocp.2014.04.033. Epub 2014 Sep 8. Review.	:25199859	Conventional	Newcastle-Ottawa scale	Random effects
50	Meta-analysis of nonsteroidal anti-inflammatory drug use and risk of atrial fibrillation.	Liu G	Am J Cardiol. 2014 Nov 15;114(10):1523-9. doi: 10.1016/j.amjcard.2014.08.015. Epub 2014 Aug 27. Review.	:25260945	Conventional	Newcastle-Ottawa Scale	Random effects
51	Sex differences in short-term and long-term all-cause mortality among patients with ST-segment elevation myocardial infarction treated by primary percutaneous intervention: a meta-analysis.	Pancholy SB	JAMA Intern Med. 2014 Nov;174(11):1822-30. doi: 10.1001/jamainternmed.2014.4762. Review.	:25265319	Conventional	Newcastle-Ottawa Scale	Random effects
52	Impact of an invasive strategy on 5 years outcome in men and women with non-ST-segment elevation acute coronary syndromes.	Alfredsson J	Am Heart J. 2014 Oct;168(4):522-9. doi: 10.1016/j.ahj.2014.06.025. Epub 2014 Jul 13. Review.	:25262262	Individual patient-level data		
53	Industry sponsorship bias in research findings: a network meta-analysis of LDL cholesterol reduction in randomised trials of statins.	Naci H	BMJ. 2014 Oct 3;349:g5741. doi: 10.1136/bmj.g5741.	:25281681 |PMCID:PMC4184241	Bayesian; Network	Cochrane Risk of Bias Tool	Fixed effect; Random effects
54	Effects of cobalt-chromium everolimus eluting stents or bare metal stent on fatal and non-fatal cardiovascular events: patient level meta-analysis.	Valgimigli M	BMJ. 2014 Nov 4;349:g6427. doi: 10.1136/bmj.g6427.	:25378023 |PMCID:PMC4220288	Individual patient-level data	Cochrane Risk of Bias Tool	Random effects
55	Anticoagulant therapy during primary percutaneous coronary intervention for acute myocardial infarction: a meta-analysis of randomized trials in the era of stents and P2Y12 inhibitors.	Bangalore S	BMJ. 2014 Nov 11;349:g6419. doi: 10.1136/bmj.g6419.	:25389143 |PMCID:PMC4227311	Network; Conventional	Cochrane Risk of Bias Tool	Fixed effect; Random effects
56	Independent influence of overweight and obesity on the regression of left ventricular hypertrophy in hypertensive patients: a meta-analysis.	Zhang K	Medicine (Baltimore). 2014 Nov;93(25):e130. doi: 10.1097/MD.0000000000000130.	:25437025 |PMCID:PMC4616382	Conventional	Modified Jadad Scale	Fixed effect; Random effects
57	A meta-analysis of mortality and major adverse cardiovascular and cerebrovascular events in patients undergoing transfemoral versus transapical transcatheter aortic valve	Panchal HB	Am J Cardiol. 2014 Dec 15;114(12):1882-90. doi: 10.1016/j.amjcard.2014.09.029. Epub 2014 Sep 28. Review.	:25438917	Conventional	Newcastle-Ottawa Scale	Fixed effect; Random effects

	implantation using edwards valve for severe aortic stenosis.						
58	Effects of blood pressure lowering on cardiovascular risk according to baseline body-mass index: a meta-analysis of randomised trials.	Blood Pressure Lowering Treatment Trialists' Colla	Lancet. 2015 Mar 7;385(9971):867-74. doi: 10.1016/S0140-6736(14)61171-5. Epub 2014 Nov 4.	:25468168	Conventional	Cochrane Risk of Bias Tool	Random effects
59	Extended duration dual antiplatelet therapy and mortality: a systematic review and meta-analysis.	Elmariah S	Lancet. 2015 Feb 28;385(9970):792-8. doi: 10.1016/S0140-6736(14)62052-3. Epub 2014 Nov 16. Review. Erratum in: Lancet. 2015 May 9;385(9980):1834.	:25467565 |PMCID:P MC438669 0	Bayesian	Cochrane Risk of Bias Tool	Random effects
60	Effects of blood pressure reduction in mild hypertension: a systematic review and meta-analysis.	Sundström J	Ann Intern Med. 2015 Feb 3;162(3):184-91. doi: 10.7326/M14-0773. Review.	:25531552	Individual patient-level data		Random effects
61	Obstructive sleep apnea and serious adverse outcomes in patients with cardiovascular or cerebrovascular disease: a PRISMA-compliant systematic review and meta-analysis.	Xie W	Medicine (Baltimore). 2014 Dec;93(29):e336. doi: 10.1097/MD.0000000000000336. Review.	:25546682 |PMCID:P MC460261 4	Conventional	Used PRISMA statement as the basis of a modified scoring system.	Random effects
62	Efficacy and safety of LDL-lowering therapy among men and women: meta-analysis of individual data from 174,000 participants in 27 randomised trials.	Cholesterol Treatment Trialists' (CTT) Collaborati	Lancet. 2015 Apr 11;385(9976):1397-405. doi: 10.1016/S0140-6736(14)61368-4. Epub 2015 Jan 9.	:25579834	Individual patient-level data		
63	Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis.	Biswas A	Ann Intern Med. 2015 Jan 20;162(2):123-32. doi: 10.7326/M14-1651. Review. Erratum in: Ann Intern Med. 2015 Sep 1;163(5):400.	:25599350	Conventional	Proper et al quality assessment tool	Fixed effect; Random effects
64	Meta-analysis appraising high maintenance dose clopidogrel in patients who underwent percutaneous coronary intervention with and without high on-clopidogrel platelet reactivity.	Ma W	Am J Cardiol. 2015 Mar 1;115(5):592-601. doi: 10.1016/j.amjcard.2014.12.013. Epub 2014 Dec 18. Review.	:25613665	Conventional	Cochrane Risk of Bias Tool	Fixed effect; Random effects
65	Effect of gender after transcatheter aortic valve implantation: a meta-analysis.	Conrotto F	Ann Thorac Surg. 2015 Mar;99(3):809-16. doi: 10.1016/j.athoracsur.2014.09.089. Epub 2015 Jan 27.	:25633460	Conventional	Modified MOOSE criteria applied to individual studies	Random effects
66	Blood pressure lowering in type 2 diabetes: a systematic review and meta-analysis.	Emdin CA	JAMA. 2015 Feb 10;313(6):603-15. doi: 10.1001/jama.2014.18574. Review.	:25668264	Conventional	Cochrane Risk of Bias Tool	Fixed effect

67	Diet quality as assessed by the Healthy Eating Index, the Alternate Healthy Eating Index, the Dietary Approaches to Stop Hypertension score, and health outcomes: a systematic review and meta-analysis of cohort studies.	Schwingshackl L	J Acad Nutr Diet. 2015 May;115(5):780-800.e5. doi: 10.1016/j.jand.2014.12.009. Epub 2015 Feb 11. Review.	:25680825	Conventional	Newcastle-Ottawa Scale	Random effects
68	Effect of Vitamin D Supplementation on Blood Pressure: A Systematic Review and Meta-analysis Incorporating Individual Patient Data.	Beveridge LA	JAMA Intern Med. 2015 May;175(5):745-54. doi: 10.1001/jamainternmed.2015.0237. Review.	:25775274	Individual patient-level data	Delphi Checklist	Random effects
69	Meta-analysis of the relation of body mass index to all-cause and cardiovascular mortality and hospitalization in patients with chronic heart failure.	Sharma A	Am J Cardiol. 2015 May 15;115(10):1428-34. doi: 10.1016/j.amjcard.2015.02.024. Epub 2015 Feb 18. Review.	:25772740	Conventional		Random effects
70	Mortality in patients treated with extended duration dual antiplatelet therapy after drug-eluting stent implantation: a pairwise and Bayesian network meta-analysis of randomised trials.	Palmerini T	Lancet. 2015 Jun 13;385(9985):2371-82. doi: 10.1016/S0140-6736(15)60263-X. Epub 2015 Mar 14. Review.	:25777667	Bayesian; Network		Fixed effect; Random effects
71	Nut consumption on all-cause, cardiovascular, and cancer mortality risk: a systematic review and meta-analysis of epidemiologic studies.	Grosso G	Am J Clin Nutr. 2015 Apr;101(4):783-93. doi: 10.3945/ajcn.114.099515. Epub 2015 Feb 4. Review.	:25833976	Conventional	Newcastle-Ottawa scale	Random effects
72	Digital health interventions for the prevention of cardiovascular disease: a systematic review and meta-analysis.	Widmer RJ	Mayo Clin Proc. 2015 Apr;90(4):469-80. doi: 10.1016/j.mayocp.2014.12.026. Review.	:25841251|PMCID:PMC4551455	Conventional	Newcastle-Ottawa Scale Modified Cochrane Risk of Bias Tool	Random effects
73	Optimal duration of dual antiplatelet therapy after percutaneous coronary intervention with drug eluting stents: meta-analysis of randomised controlled trials.	Navarese EP	BMJ. 2015 Apr 16;350:h1618. doi: 10.1136/bmj.h1618.	:25883067|PMCID:PMC4410620	Conventional	Cochrane Risk of Bias Tool	Fixed effect
74	Impact of smoking and smoking cessation on cardiovascular events and mortality among older adults: meta-analysis of individual participant data from prospective cohort studies of the CHANCES consortium.	Mons U	BMJ. 2015 Apr 20;350:h1551. doi: 10.1136/bmj.h1551.	:25896935|PMCID:PMC4413837	Individual patient-level data		Fixed effect; Random effects
75	Effects of Proprotein Convertase Subtilisin/Kexin Type 9 Antibodies in Adults With Hypercholesterolemia: A Systematic Review and Meta-analysis.	Navarese EP	Ann Intern Med. 2015 Jul 7;163(1):40-51. doi: 10.7326/M14-2957. Review.	:25915661	Conventional		Fixed effect; Random effects
76	The Volume-Outcome Relationship in Critical Care: A Systematic Review and Meta-analysis.	Nguyen YL	Chest. 2015 Jul;148(1):79-92. doi: 10.1378/chest.14-2195. Review.	:25927593|PMCID:P	Conventional		Random effects

			MC4493880			
77	Meta-Analysis of Large-Scale Randomized Trials to Determine the Effectiveness of Inhibition of the Renin-Angiotensin Aldosterone System in Heart Failure.	Emdin CA	Am J Cardiol. 2015 Jul 1;116(1):155-61. doi: 10.1016/j.amjcard.2015.03.052. Epub 2015 Apr 9.	:25937349	Conventional	Cochrane Risk of Bias Tool	Fixed effect
78	Remote Monitoring of Implantable Cardioverter-Defibrillators: A Systematic Review and Meta-Analysis of Clinical Outcomes.	Parthiban N	J Am Coll Cardiol. 2015 Jun 23;65(24):2591-600. doi: 10.1016/j.jacc.2015.04.029. Epub 2015 May 13. Review.	:25983009	Conventional	CONSORT statement	Fixed effect; Random effects
79	Comparative efficacy and safety of blood pressure-lowering agents in adults with diabetes and kidney disease: a network meta-analysis.	Palmer SC	Lancet. 2015 May 23;385(9982):2047-56. doi: 10.1016/S0140-6736(14)62459-4. Review.	:26009228	Network		Random effects
80	Complete revascularisation in ST-elevation myocardial infarction and multivessel disease: meta-analysis of randomised controlled trials.	Kowalewski M	Heart. 2015 Aug;101(16):1309-17. doi: 10.1136/heartjnl-2014-307293. Epub 2015 Jun 2.	:26037102	Conventional	Cochrane Risk of Bias Tool	Fixed effect; Random effects
81	Habitual chocolate consumption and risk of cardiovascular disease among healthy men and women.	Kwok CS	Heart. 2015 Aug;101(16):1279-87. doi: 10.1136/heartjnl-2014-307050. Epub 2015 Jun 15.	:26076934	Conventional		Random effects
82	Left Atrial Appendage Closure as an Alternative to Warfarin for Stroke Prevention in Atrial Fibrillation: A Patient-Level Meta-Analysis.	Holmes DR Jr	J Am Coll Cardiol. 2015 Jun 23;65(24):2614-23. doi: 10.1016/j.jacc.2015.04.025.	:26088300	Individual patient-level data	Newcastle-Ottawa Scale	Random effects