CLINICAL ORTHOPAEDICS AND RELATED RESEARCH
Number 455, pp. 23–29
© 2007 Lippincott Williams & Wilkins

How to Write a Systematic Review

Rick W. Wright, MD[*]; Richard A. Brand, MD[†]; Warren Dunn, MD[‡,§]; and Kurt P. Spindler, MD[‡]

Downloaded from http://journals.lww.com/clinorthop by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8K2+Ya6H515kE= on 11/02/2021

Evidence-based medicine (EBM) is the combination of the best available research evidence with clinical experience and patient needs. The concept of EBM as a part of clinical decision making has become increasingly popular over the last decade. In the hierarchy of studies meta-analysis and systematic reviews occupy the highest levels. A systematic review of a clinical question can be performed by following a relatively standard form. These techniques as described here can be performed without formal training. Systematic reviews conducted in this fashion can be used as a higher form of current concepts or as review articles and replace the traditional expert opinion narrative review.

Systematic reviews as well as meta-analyses of appropriate studies can be the best form of evidence available for clinicians. Evidence-based medicine (EBM) uses the best available research evidence along with clinical experience and patient needs and expectations.[40] According to Sackett et al,[40] the current era of recognizing these ideas and assigning the term EBM to the concepts was begun by Gordon Guyatt and a group he led at McMaster University in 1992.[15] By 1998 over 1000 publications addressed the topic. A variety of factors contributed to the increased importance of EBM: (1) physicians have a daily need for relevant information regarding diagnosis, treatment, and prognosis; (2) the traditional information sources are frequently out of date, incorrect, or overwhelming in their volume; (3) the divergence in increased clinical experience from decreased scientific study knowledge over time; and (4) time constraints with a few minutes per week for reading and review. A few recent developments have helped to overcome these barriers including: (1) new strategies for tracking and evaluating evidence; (2) evidence-based journals; (3) technological advancement for searching and acquiring the information in seconds; and (4) systematic focused reviews of healthcare studies.[40]

Historically, expert opinion has been presented in narrative reviews which are not evidence-based, and, consequently have limitations.[29,38] Unsystematic narrative reviews are more likely to include only research selected by the authors, thereby introducing bias; hence, they frequently lag behind and contradict available evidence.[2,36] However, a systematic review, as defined by Cook et al,[4] is "the application of scientific strategies that limit bias by the systematic assembly, critical appraisal and synthesis of all relevant studies on a specific topic." Systematic reviews are labor intensive and require expertise in the subject matter and review methods. The latter is particularly important because when studies are sufficiently similar a metaanalysis, which involves the statistical pooling of data from individual studies, may be appropriate. When the results of several studies seem similar, a metaanalysis can yield a more precise (narrower confidence intervals) overall estimate of the treatment effect. If studies seem sufficiently comparable and reviewers are considering pooling the results, the next step is to determine heterogeneity of the data. Determining heterogeneity requires a biostatistician or metaanalyst and a clinician with good clinical sense because, statistical tests for heterogeneity notwithstanding, ascertaining clinical heterogeneity (eg, differences in study populations) requires a clinician familiar with the subject matter.

Heterogeneity is a double-edged sword; it improves external validity at the cost of internal validity. In other words, using very narrow inclusion criteria can create more homogenous data, but at some point this process will exclude patients with certain characteristics or exposures, making data less generalizable. Hence, it can be very inappropriate to pool dissimilar studies in a metaanalysis, but it is never inappropriate to undertake a systematic

From the [*]Department of Orthopaedic Surgery, Barnes-Jewish Hospital at Washington University School of Medicine, St. Louis, MO; [†]Clinical Orthopaedics and Related Research, Philadelphia, PA; the [‡]Department of Orthopaedic Surgery and Rehabilitation, Vanderbilt University, Nashville, TN; and the [§]Division of General Internal Medicine and Public Health, Center for Health Services Research, Vanderbilt University, Nashville, TN.
Each author certifies that he or she has no commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.
Correspondence to: Rick W. Wright, MD, 1 Barnes-Jewish Plaza, Suite 11300, St. Louis, MO 63110. Phone: 314-747-2813; Fax: 314-747-2643; E-mail: wrightr@wudosis.wustl.edu.
DOI: 10.1097/BLO.0b013e31802c9098

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

review. To that end, if studies are dissimilar, precluding a metaanalysis, a descriptive summary of the studies in a systematic review should be performed.

Reviewers often narrow inclusion criteria to deal with heterogeneity by including only those studies reporting a particular outcome, or by limiting the review to specific study designs.[37] The disadvantage of this approach is it biases the review against potentially valuable studies not reporting an outcome in a specific manner.[18] Frequently, the studies meeting inclusion criteria may represent heterogeneous studies which should not be combined for statistical evaluation.[11,17,25,50] In that setting, a systematic review of available data presented in qualitative form following clearly defined methods allows orthopaedists some freedom in interpreting the best evidence.

Systematic reviews and metaanalyses of a group of Level 1 and 2 studies are the highest level of evidence (Appendix 1). Reviews of lower level studies can be performed but are thus lower levels of evidence. The steps of performing a systematic review are noted in Appendix 2.

The steps of performing a systematic review are reasonably straightforward and can be performed by many researchers (Appendix 2). The goal of our overview is to summarize the critical steps in performing a systematic review. While such overviews have been previously published (eg, Montori et al[36]), we believe this will allow readers to better understand these reviews and feel comfortable in potentially performing systematic reviews in the future. As systematic reviews become more commonly understood and accepted in orthopaedics, we believe they will ultimately be the format expected for the reviews and current concepts in our journals.

Steps in Writing a Systematic Review

Research Question

The first step in performing a systematic review is to formulate a primary research question as part of the research protocol.[50] The goal of developing a research protocol is to allow formulation of the questions and methods of the review before retrieving the literature. This helps minimize bias. A well-reasoned protocol and well-formulated research question increase the efficiency of the review by limiting the time and cost of identifying and obtaining relevant literature. At the outset, investigators must determine the level of evidence that will be included in the systematic review; for example, randomized clinical trials (RCTs) versus observational studies, and any method restrictions, such as allocation concealment. Appropriate questions to be addressed include: (1) phenomena associated with disease or interventions, (2) disease or condition frequency, (3) diagnostic accuracy, (4) disease etiology and/or risk factors, (5) prognosis, and (6) intervention effects.[16] The aims of a systematic review can be varied and include: (1) clarifying the relative strengths and weaknesses of the literature on the question, (2) summarizing a large amount of literature, (3) resolving literature conflicts, (4) evaluating the need for a large clinical trial, (5) avoiding a redundant unnecessary trial, (6) increasing the statistical power of smaller studies, (7) improving the precision or identify a smaller treatment effect, and (8) improving the generalizability of treatment outcomes.[27]

The focus of the question is an important issue. If the question is too narrow then very few studies may be identified and the generalizability to any other populations may be limited. If the question remains too broad it can be difficult to reach conclusions applicable to any single population. A well-formulated research question usually contains four parts[17] and is known by the acronym PICO: population, patient groups studied; intervention, treatment, test, or exposure for the population; comparison, alternative intervention or control; and outcome, results of the interventions.

Research Protocol

Once the research question is formulated, the research protocol is developed. The methods for literature searching, screening, data extraction, and analysis should be contained in a written document to minimize bias before starting the literature search. Strict inclusion and exclusion criteria for studies should be determined. The focus of this step is experimental design. The research question influences experimental designs considered for inclusion, for instance, RCT, experimental studies without randomization, observational studies with control groups (cohort and case-control), and observational studies without control groups (cross-sectional and case series).[25] Although method is important, the quality of a systematic review depends on the quality of the studies appraised. It can be difficult to reach meaningful conclusions from reviews of low-level evidence, and thus, systematic reviews are commonly limited to high-level evidence (Level I or II) studies (RCTs). If a research question necessitates the inclusion of low-level evidence, Level III (retrospective cohorts) or Level IV (case series), then the systematic review is likewise low-level evidence. Such reviews can be important preliminary studies, and may identify incidence of results and areas for future research (RCTs).

Literature Search

After developing the research question and protocol, a literature search commences (Appendix 3). Medline and Embase bibliographic databases have made this step more straightforward.[11,17,25,40,50] Medline includes 10 million references to journal articles since 1966. The majority of the journals referenced are published in the United States.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

EMBASE contains 8 million references to journal articles published since 1974. The EMBASE data provide better coverage of European journals. Medline can be accessed online free through PubMed. EMBASE requires a subscription to which most university or hospital libraries have access. Both databases should be searched because there is only 34% overlap of journals between the two databases. A simple Medline search is inadequate,[5] and depending on the subject of the research question, only 30% to 80% of RCTs are identified.[11] The Cochrane Collaboration has established the Cochrane Controlled Trials Register, which contains 250,000 records of controlled trials. University or hospital libraries typically have access to the Cochrane Library by CD-ROM or online. The terms searched for in these databases should be included in the manuscript as part of the search strategy.

After identifying studies using the electronic databases, the bibliographies should be reviewed to identify additional relevant studies. The bibliographies of the articles identified by this previous step should then also be reviewed. Bibliographies of the review articles in the field should be searched. Hand searching of pertinent journals for the question should also be undertaken. This is particularly important for the previous 6 months, when journal articles may not yet to be contained in electronic databases.

Publication bias, discussed later, remains an issue for systematic reviews.[7,9,21,47–49] In general, this bias exists when studies with only positive or substantial differences are published. Thus, unpublished studies (particularly those with substantial negative or inconclusive results) may contain data that would affect the overall conclusions of a systematic review. Easterbrook et al[6] found no substantial differences in quality between published and unpublished clinical studies. In one study metaanalyses limited to published trials overestimated an effect by 12% compared with those including published and grey literature (ie, unpublished, non-peer–reviewed journals, theses, industry data, hard to find articles).[30] In addition to their lack of meaningful peer-review, abstracts from meetings typically do not include enough data for a thorough review.

The goal of the literature search is to be exhaustive enough to develop a comprehensive list of potentially relevant studies. All of the studies included in the systematic review will come from this list.[50] Before finalizing the search, it is important to screen the studies to remove any duplicate studies by entering them into a computer-based referenced management system (ie, EndNote, ProCite, or Reference Manager). This also makes it easy to provide a list of initial references if a reader or journal editor requests them. A minimum of two reviewers performs a first-stage screening of titles and abstracts based on the research question and its study design, population, intervention, and outcome to be studied. Based on the initial screening, selected full-text articles are obtained for the second-stage screening. Two reviewers minimize the introduction of bias by either reviewer. Any study identified by either reviewer should be included. Using the full text a second-stage screening is performed by at least two reviewers. The studies selected are then submitted for data extraction.

Publication bias obviously affects studies found in a literature search.[1,7,9–13,16,20,21,24,34,47–49] This bias can distort the conclusions of the systematic review. Ultimate publication of studies, unfortunately, is not independent of the results. Positive studies indicating the effectiveness of a treatment are more likely to be published. An English-language bias exists where positive findings are more likely to be published in English.[13] Hence, reviews that limit articles to the English language may introduce bias. Large studies are commonly published more than once and are more likely to be cited in the bibliographies of other studies. Until recently, some editors have stated in their instructions to authors that studies not demonstrating statistically significant differences would not be published. Many of these problems are difficult to control in a systematic review, but at least language bias can be minimized by searching for foreign publications. This can add to the complexity and cost of a review by requiring translation of studies the authors are unable to read. If practical, these studies should be included. Fortunately, the effect of language bias minimally impacts the conclusions of systematic reviews.

Data Extraction

A standardized form (paper or electronic) assists in the task of data extraction.[11,17,50] The electronic form offers the advantage of simultaneous data retrieval and data entry in one step. Then, any future use of the data becomes easier. The forms should be carefully designed and piloted on a few studies before incorporation for the entire review. The data collection forms may change between different systematic reviews. Some generic items should be contained in all systematic review data collection forms (Appendix 4). Depending on the particular systematic review, more specific data collection items may need to be extracted for appropriate review, including specific items regarding population studied, intervention used, and outcomes measured. An example of a manuscript worksheet for data extraction is contained in an article by Spindler et al.[45] After data extraction, final inclusion and exclusion decisions are made regarding the manuscripts. Exclusions made at this step should be recorded, including the reason for exclusion for future reference if readers, journal editors, or reviewers desire the information. Data extraction

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

should be performed by two independent reviewers and any differences reconciled by mutual agreement.

Quality Appraisal

The key step in systematic review is quality appraisal of the included studies.[1,3,14,19,22,23,26,28,31,32,35,39,42] Several quality scales and checklists have been developed to assist in this process.[3,14,19,39,42] However, overall scores may not provide adequate information regarding individual strengths and weaknesses of the studies.[11] A study may have a raw score on a scale which indicates a quality study but contains a fatal flaw in methods.[50] All of the scales are arbitrarily subjective in the relative values assigned for different items. A study evaluating the use of these scales on a set of manuscripts reported different scales have developed widely disparate scores.[23] Therefore, most researchers agree a checklist of necessary elements for a quality study represents a more reasonable approach to quality appraisal. The items on the checklist missing in an individual study can then be presented in a qualitative manner in the systematic review. While many checklists exist, the basic format resembles the items contained in the CONSORT Statement: Revised Recommendations for Improving the Quality of Reports of Parallel-Group Randomized Trials[35] (Appendix 5).

A minimum of two independent reviewers should assess the quality of the study. Differences can be reconciled by mutual agreement or by a third reviewer. Some reviewers have suggested quality appraisal of studies should be performed in a blinded fashion by blacking out journal name, study title, authors, and institutions. This involves a substantial amount of work for the reviewers and it has been suggested the ultimate results of the systematic review are not affected by this blinding.[22]

Quality remains a difficult concept to define. However, study quality at a minimum includes internal and external validity.[11,25,50] Internal validity refers to the minimization of method error or bias in a study. External validity refers to the generalizability of the conclusions of a trial to other populations. A clinical trial may have high internal validity without applicability to other populations, but a trial without internal validity has minimal external validity. The four main biases affecting method quality include selection, performance, detection, and attrition bias.[11,25,50]

Selection bias refers to problems in the randomization process.[25,50] The allocation sequence must be unpredictable and blinded to the investigators who enroll patients. Examples of adequate allocation sequences include computer-generated random numbers, tables of random numbers, drawing lots, or envelopes. One potential problem with envelope randomization occurs when the investigator can see through the envelope. Coin tossing, shuffling cards, and rolling dice, while random, are other methods with possible allocation problems. Examples of inadequate generation of allocation include case record number, date of birth, date of admission, and alternating patients. Failure of adequate allocation can result in different populations exposed to the interventions and invalid conclusions.

Performance bias becomes an issue if additional treatments or interventions are provided to one of the groups.[11,25] Blinding of patients and treatment providers helps prevent this bias and minimizes the placebo response between the two groups.

Detection bias occurs if the investigators are influenced by the allocation sequence in assessing outcomes. This is minimized by blinding patients and investigators (including other healthcare personnel such as radiologists) measuring outcomes or administering outcome instruments. Trials inadequately reporting details regarding concealment of treatment allocation overestimate the effect of the intervention by 30% to 40% compared with trials adequately describing allocation concealment.[33,41]

Attrition bias refers to the exclusion of patients or losses to followup that occur after treatment allocation.[11,17,25,40,50] All randomized patients should be included in the analysis and kept in their original treatment groups regardless of ultimate treatment. This is the intention-to-treat principle, and it minimizes selection bias. In large clinical trials some patients will be lost to followup and final outcome cannot be assessed. Authors must discuss the number of patients lost to followup and any effect this may have on the results. A checklist to assess for these four important biases is important to data extraction and should be mentioned in any qualitative presentation of studies (Appendix 6).

Data Analysis and Results

After including and excluding studies based on the quality appraisal, data analysis and results of the studies should be undertaken.[11,17,25,32,50] The initial step for this process involves a simple descriptive evaluation of each study, commonly presented in tabular format. Tables should include the population under study, the interventions, and outcomes. Methods and biases can also be included. The decisions about items to include in the description relate back to the research question. Review of such tables can help determine if results from different studies can be pooled and subjected to a metaanalysis. When indicated, a metaanalysis can decrease random errors found in isolated studies. Again, metaanalysis is not always indicated nor feasible because of clinical heterogeneity between studies with regards to populations, interventions, or form of outcome assessment.[8,10,28,37,43,44,46] In addition, method heterogeneity in study design and quality affect the ability to perform a metaanalysis. The combination of poor-quality studies with high-quality studies will not increase the va-

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

lidity of the conclusions and, in essence, lowers the level of evidence of the review. Statistical methods and programs exist for the evaluation of heterogeneity using metaregression and for the performance of metaanalysis.[17,25,40,50] Involvement of a statistician is critical; when metaanalysis is anticipated a statistician should be involved early development of the research protocol.

Interpretation of Results

Most of this information can be presented in the data analysis and results table in the manuscript. The strengths and weaknesses of the included studies must be discussed. Metaanalysis can make interpreting the effects of the intervention easy to present and conclusions relatively straightforward. When study heterogeneity precludes metaanalysis, the authors of the systematic review need to summarize the findings based on the strength of the individual studies and reach conclusions if indicated.

The goal of this section of a systematic review is to make conclusions based on best available scientific evidence to improve clinical decision making. Frequently, the number of studies, population size, or study quality makes this difficult, and the authors should make recommendations regarding future studies. These recommendations may include study design, methods, sample size, and quality issues necessary to adequately power a future study. Systematic reviews can improve patient care by summarizing areas which have been adequately investigated and identify deficient areas to focus future research efforts and resources.

DISCUSSION

Some believe systematic reviews represent less effort than a primary clinical study. As can be seen by this review, a quality systematic review requires substantial preparation and planning. After adequate development of the research question and protocol, a considerable amount of effort is required to search the literature, appraise the study quality, and reach thoughtful, appropriate conclusions.

Systematic reviews can suffer from a variety of weaknesses during their preparation. A less than thorough literature search may miss important studies, which may affect conclusions. Frequently, orthopaedic clinical issues have not been addressed or do not lend themselves to high Level of Evidence studies. The level of conclusions reached, however, cannot exceed the level of the studies reviewed. A critical search for biases is required to adequately assess the studies. An understanding of the issues critical to studies of a particular topic is important to determine those potential biases critical to the conclusions of a study.

Despite these limitations systematic reviews can add substantially to our available evidence for clinical decision making. These deductions often represent information previously unavailable or unattainable because of study complexity or size required to answer the research question. Using the steps described in this article, a researcher can perform a systematic review even if he or she is not formally trained in the methods. This systematic approach to a critical review of all the available evidence will elevate the standards previously acceptable for nonsystematic narrative expert opinion reviews.

References

1. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gotzsche PC, Lang T. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*. 2001;134:663–694.
2. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts: treatments for myocardial infarction. *JAMA*. 1992;268:240–248.
3. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D, Hill S, Jaeschke R, Leng G, Liberati A, Magrini N, Mason J, Middleton P, Mrukowicz J, O'Connell D, Oxman AD, Phillips B, Schunemann HJ, Edejer TT, Varonen H, Vist GE, Williams JW Jr, Zaza S. Grading quality of evidence and strength of recommendations. *BMJ*. 2004;328:1490.
4. Cook DJ, Sackett DL, Spitzer WO. Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam Consultation on Meta-Analysis. *J Clin Epidemiol*. 1995;48:167–171.
5. Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ*. 1994;309:1286–1291.
6. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet*. 1991;337:867–872.
7. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315: 629–634.
8. Egger M, Schneider M, Smith GD. Meta-analysis spurious precision? Meta-analysis of observational studies. *BMJ*. 1998;316: 140–144.
9. Egger M, Smith GD. Bias in location and selection of studies. *BMJ*. 1998;316:61–66.
10. Egger M, Smith GD. Meta-analysis: unresolved issues and future developments. *BMJ*. 1998;316:221–225.
11. Egger M, Smith GD, Altman DG. *Systematic Reviews in Health Care: Meta-Analysis in Context*, 2nd ed. London: BMJ Books; 2001.
12. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. *BMJ*. 1997;315:1533–1537.
13. Egger M, Zellweger-Zahner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. *Lancet*. 1997;350:326–329.
14. Emerson JD, Burdick E, Hoaglin DC, Mosteller F, Chalmers TC. An empirical study of the possible relation of treatment differences to quality scores in controlled randomized clinical trials. *Control Clin Trials*. 1990;11:339–352.
15. Evidence-Based Medicine Working Group. Evidence-based medicine: a new approach to teaching the practice of medicine: Evidence-Based Medicine Working Group. *JAMA*. 1992;4;268:2420-2425.
16. Felson DT. Bias in meta-analytic research. *J Clin Epidemiol*. 1992; 45:885–892.
17. Glasziou P, Irwig L, Bain C, Colditz G. *Systematic Reviews in Health Care: A Practical Guide*. Cambridge, UK: Cambridge University Press; 2001.
18. Gotzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Control Clin Trials*. 1989;10:31–56.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

19. Huwiler-Muntener K, Juni P, Junker C, Egger M. Quality of reporting of randomized trials as a measure of methodologic quality. *JAMA*. 2002;287:2801–2804.
20. Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M, Moher D. Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals. *JAMA*. 1998;280:278–280.
21. Jadad AR, Moher D, Klassen TP. Guides for reading and interpreting systematic reviews: II: how did the authors find the studies and assess their quality? *Arch Pediatr Adolesc Med*. 1998;152:812–817.
22. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuay HJ. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*. 1996;17:1–12.
23. Juni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *BMJ*. 2001;323: 42–46.
24. Juni 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.
25. Khan KS, Kunz R, Kleijnen J, Antes G. *Systematic Reviews to Support Evidence-based Medicine: How to Review and Apply Findings of Healthcare Research*. London: Royal Society of Medicine Press; 2003.
26. Klassen TP, Jadad AR, Moher D. Guides for reading and interpreting systematic reviews: I: getting started. *Arch Pediatr Adolesc Med*. 1998;152:700–704.
27. Lang T, Secic M. *How to Report Statistics in Medicine*. Philadelphia: American College of Physicians; 1997.
28. Liberati A. "Meta-analysis: statistical alchemy for the 21st century": discussion: a plea for a more balanced view of meta-analysis and systematic overviews of the effect of health care interventions. *J Clin Epidemiol*. 1995;48:81–86.
29. McAlister FA, Clark HD, van Walraven C, Straus SE, Lawson FM, Moher D, Mulrow CD. The medical review article revisited: has the science improved? *Ann Intern Med*. 1999;131:947–951.
30. McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *Lancet*. 2000;356:1228–1231.
31. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement: QUOROM Group. *Br J Surg*. 2000;87:1448–1454.
32. Moher D, Jadad AR, Klassen TP. Guides for reading and interpreting systematic reviews: III: how did the authors synthesize the data and make their conclusions? *Arch Pediatr Adolesc Med*. 1998;152: 915–920.
33. Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, Tugwell P, Klassen TP. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet*. 1998;352:609–613.
34. Moher D, Pham B, Klassen TP, Schulz KF, Berlin JA, Jadad AR, Liberati A. What contributions do languages other than English make on the results of meta-analyses? *J Clin Epidemiol*. 2000;53: 964–972.
35. Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA*. 2001;285:1987–1991.
36. Montori VM, Swiontkowski MF, Cook DJ. Methodologic issues in systematic reviews and meta-analyses. *Clin Orthop Relat Res*. 2003; 413:43–54.
37. Mulrow C, Langhorne P, Grimshaw J. Integrating heterogeneous pieces of evidence in systematic reviews. *Ann Intern Med*. 1997; 127:989–995.
38. Mulrow CD. The medical review article: state of the science. *Ann Intern Med*. 1987;106:485–488.
39. Oxman AD. Checklists for review articles. *BMJ*. 1994;309: 648–651.
40. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. *Evidence-Based Medicine*. 2nd ed. Edinburgh, UK: Churchill Livingstone; 2000.
41. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*. 1995;273: 408–412.
42. Schunemann HJ, Best D, Vist G, Oxman AD. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. *CMAJ*. 2003;169:677–680.
43. Slavin RE. Best evidence synthesis: an intelligent alternative to meta-analysis. *J Clin Epidemiol*. 1995;48:9–18.
44. Smith GD, Egger M. Incommunicable knowledge? Interpreting and applying the results of clinical trials and meta-analyses. *J Clin Epidemiol*. 1998;51:289–295.
45. Spindler KP, Kuhn JE, Dunn W, Matthews CE, Harrell FE Jr, Dittus RS. Reading and reviewing the orthopaedic literature: a systematic, evidence-based medicine approach. *J Am Acad Orthop Surg*. 2005; 13:220–229.
46. Spitzer WO. The challenge of meta-analysis. *J Clin Epidemiol*. 1995;48:1–4.
47. Sterne JA, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J Clin Epidemiol*. 2001;54: 1046–1055.
48. Sterne JA, Egger M, Smith GD. Systematic reviews in health care: investigating and dealing with publication and other biases in meta-analysis. *BMJ*. 2001;323:101–105.
49. 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.
50. Torgerson C. *Systematic Reviews*. London: Continuum; 2003.

APPENDIX 1. MEDICAL EVIDENCE HIERARCHY

Meta-analysis
Systematic Review
Randomized Clinical Trials
Clinical Trials
Cohort
Case-control
Case Series
Case Reports
Expert Opinion

Appendix 1 is © 2007 and is reprinted with permission of the Journal of Bone and Joint Surgery, Inc. from Wright RW et al. Integrating Evidence-based Medicine into Clinical Practice. *J Bone Joint Surg Am.* 2007;89:199–205.

APPENDIX 2. SYSTEMATIC REVIEW STEPS

Research Question
↓
Research Protocol
↓
Literature Search
↓
Data Extraction
↓

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

Quality Assessment

↓

Data Analysis and Results

↓

Interpret Results

Appendix 2 is © 2007 and is reprinted with permission of the Journal of Bone and Joint Surgery, Inc. from Wright RW et al. Integrating Evidence-based Medicine into Clinical Practice. *J Bone Joint Surg Am.* 2007;89:199–205.

APPENDIX 3. LITERATURE SEARCH SOURCES

Sources to search for studies for systematic reviews:

Cochrane Controlled Trials Register

Medline and Embase

Other databases

Journals

Conference proceedings

Bibliographies

Unpublished and ongoing studies

APPENDIX 4. DATA EXTRACTION FORM ITEMS

Reference—*including journal, title, author, volume in page numbers*
Objective—*the study objective as stated by the authors*
Study design—*type of trial*
Population—*demographics of the participants in the study*
Intervention—*description of the intervention*
Control—*description of the control group or alternative intervention*
Outcome—*results of the intervention and how measured including statistics used*
Comments—*details regarding the study quality*

APPENDIX 5. RANDOMIZED TRIALS QUALITY CHECKLIST

Title and abstract
Introduction and background
Methods
Participants
Interventions
Objectives
Outcomes
Sample size
Randomization
Sequence generation
Allocation concealment
Implementation
Blinding
Statistical methods
Results
Participant flow
Recruitment
Baseline data
Numbers analyzed
Outcomes and estimation
Ancillary analyses
Adverse events
Comment
Interpretation
Generalizability
Overall evidence

APPENDIX 6. METHOD HIERARCHY TO ASSESS BIAS

Treatment Allocation

1. A correct blinded randomization method is described, or the study states a randomized double-blind method was used and group similarity was assessed
2. Blinded randomization stated but method not given or inadequate allocation method used
3. Randomization stated but technique not given and investigators not blinded
4. Randomization or blinding not mentioned

Bias after Allocation

1. Full and complete followup with intention to treat analysis
2. Less than 15% lost to followup with intention to treat analysis
3. No mention of withdrawals and analysis by treatment received
4. Treatment received analysis with no mention of study withdrawals or > 15% lost to followup

Blinding

1. Blinding of patient, caregiver, and investigator assessing outcome
2. Blinding of patient and caregiver or blinding of investigator assessing outcome
3. No blinding

Outcome if Unable to Blind

1. All patients assessed by standardized outcome
2. No standardized outcome measured or not mentioned in the study

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.