

ACADEMIA AND CLINIC

Systematic Reviews: Synthesis of Best Evidence for Clinical Decisions

Annals of Internal Medicine, 1 March 1997. 126:376-380.

Deborah J. Cook, MD, MSc; Cynthia D. Mulrow, MD, MSc; and R. Brian Haynes, MD, PhD

Systematic reviews can help practitioners keep abreast of the medical literature by summarizing large bodies of evidence and helping to explain differences among studies on the same question. A systematic review involves the application of scientific strategies, in ways that limit bias, to the assembly, critical appraisal, and synthesis of all relevant studies that address a specific clinical question. A meta-analysis is a type of systematic review that uses statistical methods to combine and summarize the results of several primary studies. Because the review process itself (like any other type of research) is subject to bias, a useful review requires clear reporting of information obtained using rigorous methods. Used increasingly to inform medical decision making, plan future research agendas, and establish clinical policy, systematic reviews may strengthen the link between best research evidence and optimal health care.

Ann Intern Med. 1997;126:376-380. *Annals of Internal Medicine* is published twice monthly and copyrighted © 1997 by the American College of Physicians.

From McMaster University, Hamilton, Ontario, Canada; and University of Texas Health Science Center and Veterans Affairs Hospital, San Antonio, Texas. For current author addresses, see end of text.

At the end of a long week in the office, you sink back into your chair, reflecting on some of the more memorable patients you cared for and counseled. Through gentle history taking, you discovered that urinary incontinence is the underlying cause of an elderly patient's increasing social isolation. During a careful physical examination, you detected bruising on the torso of a woman with chronic headaches and began to explore the longstanding abusive relationship between the woman and her alcoholic partner. You discontinued procainamide therapy in a 72-year-old man who had asymptomatic premature ventricular contractions after myocardial infarction. To prevent bleeding from esophageal varices, you started beta-blocker therapy in a woman with long-standing cryptogenic cirrhosis and portal hypertension. In couples' therapy, discussing the future quality of life of a middle-aged gay man with human immunodeficiency virus infection, you journeyed through emotionally intense dialogue about advance directives. You presented the risk factors for major and minor bleeding to a 39-year-old woman who was considering warfarin therapy because of recently diagnosed atrial fibrillation and valvular heart disease. You listened to, made diagnoses for, treated, advised, and comforted many patients.

Yet there were some hiccoughs in your practice along the way. You stumbled while debating the pros and cons of breast cancer screening with a healthy 48-year-old woman who has been staying current with information on the Internet. You questioned the merits of a personalized walking program suggested to you by a motivated 66-year-old man with severe claudication. Explaining that you wanted to review the best current evidence on these issues, you resolved to address your uncertainties before these patients made their next office visits, in a week's time.

Sighing deeply, you acknowledge that you have little time to read. You subscribe to three journals, which you browse months after they arrive—either when your journal stack becomes precariously high or when your guilt is sufficiently motivational. You sometimes find the conclusions of individual articles conflicting or confusing. You know that some of the decisions and suggestions you made this week, specifically your decisions about stopping procainamide therapy and starting beta-blocker therapy and your advice about bleeding risks from anticoagulant therapy, were based on the best current research evidence (1-3). On the other hand, your patients' inquiries about breast cancer screening and exercise treatment for claudication highlight your need for a concise, current, rigorous synthesis of the best available evidence on each of these topics: in brief, a systematic review (4, 5).

Incorporating Research Evidence into Clinical Decision Making

The foregoing scenario is familiar to practitioners. In a typical week, we encounter patients with diverse problems; exercise numerous clinical, interpersonal, and technical skills; and make many decisions. The factors that affect these decisions and their outcomes are complex. For instance, each patient has unique sociodemographic characteristics, cultural circumstances, and personal preferences. Each physician has unique knowledge, experiences, and values. Moreover, practitioners and their patients make decisions within the context of a rapidly changing health care system that influences the availability, accessibility, and cost of diagnostic tests and therapies (6).

Timely, useful evidence from the biomedical literature should be an integral component of clinical decision making. If one treatment has been shown to be better than another, we need to know, so that we can recommend the treatment to the appropriate patients. The worldwide effort to develop new tests and treatments, and to determine their usefulness, has never been stronger, and our patients and their families expect us to be fountains of the knowledge that results from this effort (7). Unfortunately, it is easy for current best research evidence to pass us by (8). We may lack the time, motivation, and basic skills needed to find, critically appraise, and synthesize information, all of which we must do if we are to integrate the results of original studies into our practice. Fortunately, several potent methods are emerging that can greatly enhance our ability to interpret and apply research evidence; foremost among them is the systematic review.

This article begins a series in *Annals* that will examine systematic reviews in detail and explore their many applications. Systematic reviews represent the best chance that most practitioners will have to understand and accurately apply the key signals arising from the robust and increasingly productive search for solutions to medical problems. A properly conducted systematic review faithfully summarizes the evidence from all relevant studies on the topic of interest, and it does so concisely and transparently.

What Is a Systematic Review?

Systematic reviews are scientific investigations in themselves, with pre-planned methods and an assembly of original studies as their "subjects." They synthesize the results of multiple primary investigations by using strategies that limit bias and random error (9, 10). These strategies include a comprehensive search of all potentially relevant articles and the use of explicit, reproducible criteria in the selection of articles for review. Primary research designs and study characteristics are appraised, data are synthesized, and results are interpreted.

When the results of primary studies are summarized but not statistically combined, the review may be called a qualitative systematic review. A quantitative systematic review, or meta-analysis, is a systematic review that uses statistical methods to combine the results of two or more studies. The term "overview" is sometimes used to denote a systematic review, whether quantitative or qualitative. Summaries of research that lack explicit descriptions of systematic methods are often called narrative reviews.

Review articles are one type of integrative publication; practice guidelines, economic evaluations, and clinical decision analyses are others. These other types of integrative articles often incorporate the results of systematic reviews. For example, practice guidelines are systematically developed statements intended to assist practitioners and patients with decisions about appropriate health care for or specific clinical circumstances (11). Evidence-based practice guidelines are based on systematic reviews of the literature, appropriately adapted to local circumstances and values. Economic evaluations compare both the costs and the consequences of different courses of action; the knowledge of consequences that are considered in these evaluations is often generated by systematic reviews of primary studies. Decision analyses quantify both the likelihood and the valuation of the expected outcomes associated with competing alternatives.

Differences between Systematic and Narrative Reviews

All reviews, narrative and systematic alike, are retrospective, observational research studies and are therefore subject to systematic and random error. Accordingly, the quality of a review—and thus its worth—depends on the extent to which scientific review methods have been used to minimize error and bias. This is the key feature that distinguishes traditional narrative reviews from systematic reviews (Table). If a review is prepared according to the steps outlined in the right column of the Table, it is more likely to be systematic and to provide unbiased conclusions. If review methods approximate those found in the middle column of the Table, the article is more likely to be a narrative review, and the conclusions are less likely to be based on an unbiased summary of all relevant evidence.

Systematic reviews are generated to answer specific, often narrow, clinical questions in depth. These questions can be formulated explicitly according to four variables: a specific population and setting (such as elderly outpatients), the condition of interest (for example, hypertension), an exposure to a test or treatment (such as pharmacologic management), and one or more specific outcomes (such as cardiovascular and cerebrovascular events and mortality) (12). Thus, an example of a well-formulated, clinically relevant question is, Does

pharmacologic treatment of hypertension in the elderly prevent strokes and myocardial infarctions or delay death? If the question that is driving the review is not clear from the title, abstract, or introduction, or if no methods section is included, the paper is more likely to be a narrative review than a systematic review (13).

Most narrative review articles deal with a broad range of issues related to a given topic rather than addressing a particular issue in depth (9). For example, a narrative review on diabetes (such as that which might be found in a textbook chapter) might include sections on the physiology and pathophysiology of carbohydrate, lipid, and protein metabolism; the epidemiology of and prognosis associated with diabetes; diagnostic and screening approaches; and preventive, therapeutic, rehabilitative, and palliative interventions. Thus, narrative reviews may be most useful for obtaining a broad perspective on a topic; they are less often useful in furnishing quantitative answers to specific clinical questions.

Narrative reviews are appropriate for describing the history or development of a problem and its management. Narrative reviews may better describe cutting-edge developments if research is scant or preliminary or if studies are very limited by flawed design or execution (13). They may be particularly useful for discussing data in light of underlying theory and context. Narrative reviews can draw analogies and can conceptually integrate two independent fields of research, such as cancer and the acquired immunodeficiency syndrome (13).

However, the connection between clinical recommendations and evidence in narrative reviews is often tenuous, incomplete, or worse still based on a biased citation of studies (14, 15). As a result, recommendations found in narrative reviews published in journals and textbooks often differ from recommendations found in systematic reviews. For example, narrative reviews may lag behind by more than a decade in endorsing a treatment of proven effectiveness, or they may continue to advocate a therapy long after it has been shown to be useless or harmful (16). Also, systematic reviews that incorporate quantitative techniques are more likely than narrative reviews to detect small but clinically meaningful treatment effects (17).

What Systematic Reviews Can and Cannot Do

A well-conducted systematic review is invaluable for practitioners. Many of us feel overwhelmed by the volume of medical literature and, as a result, often prefer summaries of information to publications of original investigations (18). Thus, review articles can help us keep up-to-date. High-quality systematic reviews can define the boundaries of what is known and what is not known and can help us avoid knowing less than has been proven.

It is unusual for single studies to provide definitive answers to clinical questions (19), but systematic reviews can help practitioners solve specific clinical problems. By critically examining primary studies, systematic reviews can also improve our understanding of inconsistencies among diverse pieces of research evidence. By quantitatively combining the results of several small studies, meta-analyses can create more precise, powerful, and convincing conclusions. An example of this is the recent review highlighting the beneficial effect of dietary protein restriction on the progression of diabetic and nondiabetic renal disease (20). In addition, systematic reviews of several studies may better inform us about whether findings can be applied to specific subgroups of patients (21).

Investigators need systematic reviews to summarize existing data, refine hypotheses, estimate sample sizes, and help define future research agendas. Without systematic reviews, researchers may miss promising leads or may embark on studies of questions that have been already answered. Administrators and purchasers need review articles and other integrative publications to help generate clinical policies that optimize outcomes using available resources (22).

Systematic reviews can aid, but can never replace, sound clinical reasoning. Clinicians reason about individual patients on the basis of analogy, experience, heuristics, and theory as well as research evidence (23-25). Awareness of a treatment's effectiveness does not confer knowledge about how to use that treatment in caring for individual patients. Evidence can lead to bad practice if it is applied in an uncritical or unfeeling way (25). Understanding the complex structure of medical decision making (26) requires an appreciation of the ways in which knowledge, skills, values, and research evidence are integrated in each patient-clinician encounter.

The Past, Present, and Future of Systematic Reviews

The epidemiology of review articles is changing rapidly; the number of systematic reviews published annually has increased at least 500-fold in the past decade (27, 28). It is not unusual now to find more than one systematic review addressing the same or similar therapeutic questions (29). One example is the evaluation of the effect of calcium on blood pressure, which was summarized in two recent meta-analyses published within 1 month of each other (30, 31). Although duplicate independent reviews that generate similar results may increase our confidence in those results, multiple reviews with similar results may have an opportunity cost—they may divert effort away from the development of systematic reviews in other areas where they may be needed. The situation can become more confusing when several reviews yield similar results but present different clinical recommendations; this occurred when five meta-analyses of selective decontamination of the digestive tract in critically ill patients were published within 5 years (two were published in the same issue of the same journal)

(32-36). However, resolving discord between meta-analyses can be a cooperative, informative exercise that clarifies review methods and synthesizes all currently available evidence (37).

In the past 10 years, guides have been created to aid in the critical appraisal (38, 39) and application (40) of review articles. A framework for interpreting primary research and systematic reviews by using levels of evidence and grades of recommendations (41) has been successfully employed by several groups. At the Antithrombotic Consensus Conferences of the American College of Chest Physicians, therapeutic recommendations are routinely classified according to their evidentiary base—whether they are predicated on the results of large, rigorous, randomized trials or on meta-analyses, observational studies, or expert opinion (41).

An international initiative called the Cochrane Collaboration has evolved to help prepare, maintain, and disseminate the results of systematic reviews of health care interventions (42, 43). The Cochrane Library is the first large-scale, multidisciplinary product of this collaboration. It is updated quarterly, and in January 1997 it contained 159 completed systematic reviews and 199 systematic reviews in preparation from the Cochrane Collaboration itself, a bibliography of more than 1600 other systematic review abstracts in the medical literature, a register of controlled trials that included slightly more than 110 000 trials, and a bibliography of 400 methodologic articles about the science of reviewing research (44).

What does the future hold for systematic reviews? The coverage of medical topics by systematic reviews, although expanding rapidly, is still limited. The Cochrane Collaboration is facilitating this mission but needs increasing support and participation if it is to succeed. Investigation into the science of research synthesis will increase the quality, and thus the value, of the evidence found in systematic reviews. More efficient searching methods with which to identify systematic reviews will be developed. Patients and administrators will have increasing access to reviews, which will spur practitioners to locate, appraise, and apply reviews in their practice.

Information technology can provide point-of-care access, but practitioners will need new skills to use such technology. Presentation formats will become more user-friendly for both providers and patients. Practical techniques for incorporating information from systematic reviews into clinical decision making will be created and disseminated.

While looking forward to these advances, clinicians can and should take advantage of what systematic reviews have to offer now. The *Annals* series will aid in that process by focusing on how to find, assess, use, and conduct systematic reviews for clinical, teaching, and research purposes.

Acknowledgments: We would like to acknowledge many persons for their advice and support, particularly the following.

Advisors: Gordon Guyatt (Hamilton, Ontario, Canada), R. Brian Haynes (Hamilton, Ontario, Canada), Andrew Oxman (Oslo, Norway), and David L. Sackett (Oxford, the United Kingdom).

Reviewer-Commentators: Iain Chalmers (Oxford, the United Kingdom), Robert Fletcher (Boston, Massachusetts), Larry V. Hedges (Chicago, Illinois), Andreas Laupacis (Ottawa, Ontario, Canada), C. David Naylor (North York, Ontario, Canada), and David L. Simel (Durham, North Carolina).

Clinical Reviewers: Clifton R. Cleaveland (Chattanooga, Tennessee), Paul F. Speckart (San Diego, California), and Norman J. Wilder (Anchorage, Alaska).

Local Editorial Assistants: Linn Morgan and Barbara Hill.

Grant Support: Dr. Cook is a Career Scientist of the Ontario Ministry of Health; Dr. Mulrow is Senior Research Associate at the Audie L. Murphy Memorial Veterans Hospital; and Dr. Haynes is a National Health Scientist of the National Health Research and Development Program of Health Canada.

Requests for Reprints: Deborah J. Cook, MD, MSc, Department of Medicine, St. Joseph's Hospital, 50 Charlton Avenue East, Hamilton, Ontario L8N 4A6, Canada.

Current Author Addresses: Dr. Cook: Department of Medicine, St. Joseph's Hospital, 50 Charlton Avenue East, Hamilton, Ontario L8N 4A6, Canada.

Dr. Mulrow: Audie L. Murphy Veterans Memorial Hospital, 7400 Merton Minter Boulevard, San Antonio, TX 78284.

Dr. Haynes: Health Information Research Unit, Department of Clinical Epidemiology and Biostatistics, McMaster University Medical Center, 1200 Main Street West, Hamilton, Ontario L8N 3Z5, Canada.

References

1. **Teo KK, Yusuf S, Furberg CD.** Effects of prophylactic antiarrhythmic drug therapy in acute myocardial infarction. An overview of results from randomized controlled trials. *JAMA*. 1993;270:1589-95.
2. **Pagliari L, D'Amico G, Sorensen TI, Lebrec D, Burroughs AK, Morabito A, et al.** Prevention of first bleeding in cirrhosis. A meta-analysis of randomized trials of nonsurgical treatment. *Ann Intern Med*. 1992;117:59-70.
3. **Landefeld CS, Beyth RJ.** Anticoagulant-related bleeding: clinical epidemiology, prediction, and prevention. *Am J Med*. 1993;95:315-28.
4. **Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL.** Efficacy of screening mammography. A meta-analysis. *JAMA*. 1995;273:149-54.
5. **Gardner AW, Poehlman ET.** Exercise rehabilitation programs for the treatment of claudication pain. A meta-analysis. *JAMA*. 1995;274:975-80.
6. **Sackett DL, Haynes RB.** On the need for evidence-based medicine. *Evidence-Based Medicine*. 1995;1:5-6.
7. Evidence-based medicine. A new approach to teaching the practice of medicine. *Evidence-Based Medicine Working Group*. *JAMA*. 1992;268:2420-5.
8. **Shin JH, Haynes RB, Johnston ME.** The effect of problem-based, self-directed undergraduate education on life-long learning. *Can Med Assoc J*. 1993;148:969-76.
9. **Mulrow CD.** The medical review article: state of the science. *Ann Intern Med*. 1987;106:485-88.
10. **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;48:167-71.
11. **Woolf SH.** Practice guidelines: a new reality in medicine. I. Recent developments. *Arch Intern Med*. 1990;150:1811-8.
12. **Richardson WS, Wilson MC, Nishikawa J, Hayward RS.** The well-built clinical question: a key to evidence-based decisions [Editorial]. *ACP J Club*. 1995;123:A12.
13. **Bangert-Drowns RL.** Misunderstanding meta-analysis. *Evaluation and the Health Professional*. 1995;18:304-14.
14. **Ravnskov U.** Cholesterol lowering trials in coronary heart disease: frequency of citation and outcome. *BMJ*. 1992;305:15-9.
15. **Neihouse PF, Priske SC.** Quotation accuracy in review articles. *DICP*. 1989;23:594-6.
16. **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;268:240-8.
17. **Cooper HM, Rosenthal R.** Statistical versus traditional procedures for summarizing research findings. *Psychol Bull*. 1980;87:442-9.
18. **Williamson JW, German PS, Weiss R, Skinner EA, Bowes F 3d.** Health science information management and continuing education of physicians. A survey of U.S. primary care practitioners and their opinion leaders. *Ann Intern Med*. 1989;110:151-60.
19. **Davidoff F.** Evidence-based medicine: why all the fuss? [Editorial] *Ann Intern Med*. 1995;122:727.
20. **Pedrini MT, Levey AS, Lau J, Chalmers TC, Wang PH.** The effect of dietary protein restriction on the progression of diabetic and nondiabetic renal diseases: a meta-analysis. *Ann Intern Med*. 1996;124:627-32.
21. Collaborative overview of randomised trials of antiplatelet therapy-I: Prevention of death, myocardial infarction, and stroke by prolonged antiplatelet therapy in various categories of patients. *Antiplatelet Trialists' Collaboration*. *BMJ*. 1994;308:81-106.
22. **Milne R, Hicks N.** Evidence-based purchasing. *Evidence-Based Medicine*. 1996;1:101-2.
23. **Tanenbaum SJ.** What physicians know. *N Engl J Med*. 1993;329:1268-71.
24. **McDonald CJ.** Medical heuristics: the silent adjudicators of clinical practice. *Ann Intern Med*. 1996;124(1 Pt 1):56-62.
25. **Naylor CD.** Grey zones of clinical practice: some limits to evidence-based medicine. *Lancet*. 1995;345:840-2.
26. **Dowie J.** "Evidence-based", "cost-effective" and "preference-driven" medicine: decision analysis-based medical decision making is the pre-requisite. *Journal of Health Services Research and Policy*. 1996;1:104-13.
27. **Chalmers TC, Lau J.** Meta-analytic stimulus for changes in clinical trials. *Stat Methods Med Res*. 1993;2:161-72.
28. **Chalmers I, Haynes RB.** Reporting, updating, and correcting systematic reviews of the effects of health care. In: Chalmers I, Altman DG, eds. *Systematic reviews*. London: BMJ; 1995:86-95.
29. **Jadad A, Cook DJ, Browman G.** When arbitrators disagree: resolving discordant meta-analysis. *Can Med Assoc J*. 1997; [In press].
30. **Allender PS, Cutler JA, Follmann D, Cappuccio FP, Pryer J, Elliott P.** Dietary calcium and blood pressure: a meta-analysis of randomized clinical trials. *Ann Intern Med*. 1996;124:825-31.
31. **Bucher HC, Cook RJ, Guyatt GH, Lang JD, Cook DJ, Hatala R, Hunt DL.** Effects of dietary calcium supplementation on blood pressure. A meta-analysis of randomized controlled trials. *JAMA*. 1996;275:1016-22.
32. **Vandenbroucke-Grauls CM, Vandenbroucke JP.** Effect of selective decontamination of the digestive tract on respiratory tract infections and mortality in the intensive care unit. *Lancet*. 1991;338:859-62.
33. **Hurley JC.** Prophylaxis with enteral antibiotics in ventilated patients: selective decontamination or selective cross-infection? *Antimicrobial Agents Chemother*. 1995;39:941-7.
34. Meta-analysis of randomized controlled trials of selective decontamination of the digestive tract. *Selective Decontamination of the Digestive Tract Trialists' Collaborative Group*. *Br Med J*. 1993;307:525-32.

- 35. Heyland DK, Cook DJ, Jaeschke R, Griffith L, Lee HN, Guyatt GH.** Selective decontamination of the digestive tract. An overview. *Chest*. 1994;105:1221-9.
- 36. Kollef MH.** The role of selective digestive tract decontamination on mortality and respiratory infections. A meta-analysis. *Chest*. 1994;105:1101-8.
- 37. Cook DJ, Reeve BK, Guyatt GH, Griffith LE, Heyland DK, Buckingham L, et al.** Stress ulcer prophylaxis in critically ill patients. Resolving discordant meta-analyses. *JAMA*. 1996;275:308-14.
- 38. L'Abbe KA, Detsky AS, O'Rourke K.** Meta-analysis in clinical research. *Ann Intern Med*. 1987;107:224-33.
- 39. Oxman AD, Cook DJ, Guyatt GH.** Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. *JAMA*. 1994;272:1367-71.
- 40. Guyatt GH, Sackett DL, Sinclair J, Hayward R, Cook DJ, Cook RJ.** Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. *JAMA*. 1995;274:1800-4.
- 41. Cook DJ, Guyatt GH, Laupacis A, Sackett DL, Goldberg RJ.** Clinical recommendations using levels of evidence for antithrombotic agents. *Chest*. 1995;108(4 Suppl):227S-30S.
- 42. Chalmers I.** The Cochrane Collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. *Ann N Y Acad Sci*. 1993;703:153-63.
- 43. Bero L, Rennie D.** The Cochrane Collaboration. Preparing, maintaining, and disseminating systematic reviews of the effects of health care. *JAMA*. 1995;274:1935-8.
- 44. Cochrane Collaboration.** Cochrane Library. Electronic serial publication issued quarterly by BMJ Publishing Group, London.