



EVIDENCE-BASED MEDICINE AND PRACTICE GUIDELINES: AN OVERVIEW

Steven H. Woolf, MD, MPH

From the Department of Family Practice, Medical College of Virginia - Virginia Commonwealth University, Fairfax, Virginia.

Introduction

Recent years have witnessed a growing emphasis on evidence. The volume of scientific information has increased dramatically. Thousands of randomized, controlled trials (RCTs) are published each year. Systematic reviews,¹ which summarize the evidence from multiple studies, and meta-analyses,² which quantitatively pool data, are increasingly common in general medicine journals and specialty journals. The Cochrane Collaboration, a worldwide effort to conduct systematic reviews, is continuously updating its reports (and other systematic reviews) on an electronic database, the *Cochrane Library*, which is accessible on CD-ROMs and the Internet.³

Evidence-based medicine (EBM), which seeks to directly link clinical practice and policy decisions to supporting evidence, is receiving increased attention in practice guidelines and public policy. Insurers' decisions about coverage, government decisions about health care financing, and even medicolegal judgments are guided by the results of clinical trials, systematic reviews, and pronouncements in evidence-based practice guidelines. Through its Agency for Healthcare Research and Quality, the federal government has established 12 evidence-based practice centers in institutions in the United States and Canada to perform and publish high-quality evidence reports.

Publications and electronic resources have stepped in to help

clinicians and policy makers deal with this explosion of information. Medical journals routinely publish systematic reviews and/or sections that summarize reviews in other journals. Several new journals are devoted entirely to evidence-based medicine, such as *ACP Journal Club*, *Evidence-Based Medicine*, *Best Evidence*, and *Evidence-Based Oncology*.

The Growing Emphasis on Evidence

The increased attention to data in clinical practice and health policy settings stems from three factors. First, spiraling health care costs have placed pressure on government and private health plans to scrutinize the effectiveness and appropriateness of service utilization. Second, studies since the 1970s have documented substantial geographic variations in the rate with which procedures and treatments are offered to patients with similar clinical profiles, suggesting that services are either over- or under-utilized.⁴

Studies find that services that patients should be receiving (eg, childhood immunizations, smoking cessation counseling, beta blockers after acute myocardial infarction, warfarin for nonvalvular atrial fibrillation) are not being provided in a large proportion of cases. Conversely, services that are ineffective (eg, routine preoperative testing) are performed with regularity. In other cases, services are misapplied to the wrong patient populations or are deliv-

ered incorrectly. Thrombolytic therapy is often not administered quickly enough to be effective. In other cases, frank medical errors cause up to 100,000 deaths each year in the United States.⁵

Distinguishing Features of EBM

Seasoned clinicians sometimes question the novelty of EBM, noting that medicine has always been “evidence-based.” Indeed, since the times of Ancient Greece, physicians have engaged in scientific study and have tried to keep abreast of the latest evidence. What distinguishes EBM from the traditional application of evidence in medicine, however, is the explicit linkage between policy and supporting data.

EBM emphasizes an examination of the evidence that is comprehensive, critical, and explicit. Comprehensiveness is important to ensure that all evidence is considered rather than just those studies that support a particular viewpoint or that reflect a selection bias. Critical appraisal is emphasized to examine the strengths and weakness of the study designs so that judgments about the evidence can be linked to quality. Explicitness gives transparency to the evaluation, allowing readers to understand the methods used in the analysis, the strengths of the evidence, the gaps that exist, and the rationale for practice recommendations or policies, whether evidence- or opinion-based.

Misconceptions about EBM abound.⁶ Chief among these is the presumption that EBM resists any medical practice that has not been proven in a RCT or other study. Many believe that EBM seeks to convert clinical practice into “cookbook medicine,” impeding efforts by clinicians to exercise clinical judgment and individualize care. Although this occurs, sometimes under the false guise of EBM, it is neither advocated by EBM nor defensible on rational grounds. Very little of what is done in clinical practice has been tested in controlled studies. It would therefore be unrealistic and invalid to withhold services based on this criterion. Studies such as RCTs are difficult to design and fund, often require years of follow-up to achieve results, and rely on outcome measures that may not capture the range of benefits and harms associated with interventions.^{7,8}

Evidence that is available is often of poor quality, either in terms of *internal validity* (the extent to which the data are reflective of the clinical setting in which the study was conducted) or *external validity* (the extent to which the findings can be extrapolated to other patient populations, providers, or settings). Even the best evidence can provide only “averages” for predicting outcomes in a given patient. Individual variables (eg, risk factors, past medical history, personal circumstances, provider skills, community resources) influence where a patient will fall in the bell curve that surrounds the mean.

EBM does not gloss over these considerations but insists on making them explicit. It does not insist on evidence from RCTs but does demand that the grade and quality of the evidence be carefully evaluated and stated clearly. EBM does not preclude the use of opinion or expert judgment in setting practice policy but does insist on acknowledging when this is done. It advocates disclosure of gaps in the evidence to help clarify research agendas and calls attention to design features that future studies should incorporate to address deficiencies in current data. Finally, EBM does not mask over the importance of individualized care. It encourages that the determinant factors in practice and policy decisions be evidence-based to maximize equity and effectiveness for all patients.

Evidence vs Expert Opinion

A central thesis of EBM is the putative superiority of scientific evidence over opinion. The truth is that both evidence and opinion have their limitations. As already noted, scientific evidence is absent for much of medicine and, when available, often lacks internal and external validity.

The problems with opinion, whether it be clinical judgment or the beliefs of experts, are also well documented.⁹ These include the selective use of evidence (inadvertently or consciously ignoring studies suggesting another view), biases about magnitudes of effect and appropriateness that stem from

personal experience (eg, how one was trained, a notably bad outcome in a past patient), flawed assumptions about the frequency or natural history of diseases, and external influences (eg, professional norms, business pressures, patient expectations, medicolegal concerns).¹⁰

Medical history is replete with examples of interventions that were thought to be beneficial and that received enthusiastic support from clinicians, even in the face of emerging evidence to the contrary. The tendency of physicians to assume that a treatment is beneficial — based on biological plausibility, prevailing theories about natural history, and intermediate (rather than health) outcomes — has inspired many mistakes in medical dogma. For example, cardiologists have long believed that routine antiarrhythmic therapy is beneficial or even life-saving in patients with acute myocardial infarction because of the evidence that arrhythmias are associated with increased mortality. This inference went unchallenged until the early 1990s, when clinical trials showed that antiarrhythmic therapy *increased* mortality.¹¹

However, science is not the last word, nor the only consideration, in determining what is best for patients. Although science is of paramount importance, good clinical decisions and sound public policies address other considerations, such as expert opinion and clinical experience, patient expectations and preferences, social circumstances, the time, equipment, and personnel avail-

able in the clinical setting, community support services, access to care, ethics, insurance coverage, and medicolegal risks.

Practice Guidelines

For decades, practice guidelines — official statements from organizations and agencies regarding the appropriate use of procedures and treatments — have relied on informal methods to weight these issues. Typically, a group of experts gather around a conference table and are asked to make recommendations, based on their understanding of the evidence and on their opinion. No formal methodology for reviewing evidence or arriving at consensus is followed, and whatever methods and rationale are used tend not to be recorded in the final document. Formal consensus development techniques have been used, most notably in National Institutes of Health Consensus Development Conferences¹² and the Rand appropriateness panels,¹³ but even in these approaches, the panelists still reach conclusions based on opinion.

Disfavor over these methods has grown in recent years due to concerns about vulnerability to bias and conflicts of interest.¹⁴ Many groups have turned to EBM for guideline development. Evidence-based practice guidelines feature an explicit methodology, include as their foundation a systematic review of the evidence, provide graded recommendations that are linked directly to the sup-

porting evidence, and state explicitly when recommendations are based on opinion.

Developing Evidence-Based Practice Guidelines

In general, evidence-based practice guidelines emerge from six steps that are conducted with varying intensity and in different sequences, depending on the topic.

1. Specification of Topic and Methodology

The first step is to give precision to the focus of the review, specifying the target condition, the interventions to be reviewed, relevant patient populations and clinical settings, and outcome measures of significance. The boundaries for the search are also determined, such as bibliographic databases and exclusion criteria (eg, studies published before a given date, foreign-language articles, editorials, uncontrolled studies, nonhuman studies). An evidence model often helps to clarify the linkages in the analytic framework for which evidence is sought.¹⁵

2. Systematic Review

The review of evidence follows procedures that have become standardized in recent years.¹ Three basic steps include (1) a comprehensive literature search, using explicitly documented search terms and other techniques to assure the reviewers and readers that all relevant evidence has

been gathered, (2) critical appraisal of individual studies, using explicit analytic criteria to judge internal and external validity and documenting the findings in abstraction forms and evidence tables, and (3) synthesis of results, summarizing the results in narrative text, evidence tables, or balance sheets. The last step may involve quantitative pooling of data in meta-analyses to estimate overall effect sizes, especially when individual studies lack statistical power, or in decision analytic models that predict outcomes under varied assumptions about determinant variables.

In many systematic reviews, studies are assigned a "grade," or evidence code, that reflects the position of the study in a hierarchy of evidence quality. Several coding schemes exist.¹⁶ In grading studies of the effectiveness of treatments, a common feature is to place RCTs at the top of the hierarchy, followed by observational and epidemiologic studies. Other coding schemes are appropriate for studies evaluating diagnostic tests, epidemiologic trends, and natural history.¹⁷

Reviewers examine a variety of factors to assess internal validity (eg, study population, allocation to groups, interventions, outcome measures, attrition rates, statistical measurements).¹⁸ The generalizability of the study population, intervention, and setting are considered to judge external validity. Over 20 instruments are available to grade the quality of RCTs,¹⁹ but the Jadad scale is most well known and validated.²⁰

3. Expert Opinion

Expert opinion plays a role in all practice guidelines. Even when evidence is available, subjective judgments are made in assessing the strength or generalizability of the evidence and in weighing the tradeoffs between benefits and harms. When evidence is lacking, groups differ on the extent to which they are willing to make recommendations based on opinion. A hallmark of EBM is to be explicit when opinion is used so that readers understand the basis for the recommendations and can make their own judgment about validity.

4. Public Policy Considerations

In an era of limited health care resources, guideline developers must often consider the cost effectiveness or cost utility of interventions. Other policy considerations, such as access to care, availability of qualified personnel and technology, insurance policies, and medicolegal implications are considered to varying degrees depending on the topic and panel philosophy. It is in this context that conflicts of interest among panel members become especially problematic.²¹ Some groups rigidly avoid making opinion-based recommendations, instead offering the neutral conclusion that there is insufficient evidence to make a recommendation.

5. Drafting of Document

The wording of recommendations receives great attention in practice guidelines since even the

slightest nuances of language can have serious policy implications and affect the quality of patient care. A characteristic of evidence-based guidelines is the use of letter codes (eg, "A" recommendation) or recommendation categories (eg, "standards," "guidelines," "options") to reflect how strongly the intervention is recommended. Almost always, this grading scheme reflects the strength of supporting evidence.

For interventions with complex tradeoffs or equivocal evidence, the best choice may depend on the values patients assign to potential benefits and harms. Increasingly, guidelines on such topics eschew universal recommendations and instead urge physicians to help patients assess and weigh personal preferences in a process of shared decision making.²² That process involves discussing with patients the potential benefits and harms of interventions and their likelihood, helping them decide how strongly they feel about potential outcomes, and determining from this information which option is best.

6. Peer Review

As with any scholarly document, evidence-based guidelines are typically circulated in draft form to content experts to obtain feedback on the comprehensiveness of the review and the validity of the critical appraisal. The draft is also sent to stakeholders, such as relevant professional societies and advocacy organizations, for further feedback.

Limitations of Evidence-Based Practice Guidelines

Evidence-based practice guidelines, like all guidelines, can be flawed, advocating interventions that are not in the best interest of patients.²³ Sometimes the errors stem from limitations in the science itself, such as lack of data or poor generalizability. Sometimes errors occur when panel members reach invalid conclusions in translating science into policy. Biases or conflicts of interest among panel members, often exacerbated when outspoken individuals dominate the process, can produce different recommendations than the data support. Recommendations that do not give guidance on individualization or that reduce complex decisions into simplistic algorithms may be overly rigid and may result in more harm than good.

Practice guidelines can have adverse implications for clinicians, especially if they are rigidly enforced by payers, managers, or malpractice courts. They can have adverse policy implications for society if they increase the costs of care, decrease equity, or divert resources from more effective health care interventions.

A fundamental limitation of practice guidelines is that they often do not change practice behavior. Most studies indicate that passive dissemination of guidelines, such as publishing them in a medical journal, is ineffective in changing behavior.²⁴ Guidelines have been shown to be effective in changing practice patterns when

they are accompanied by active implementation strategies, such as standing orders, reminder systems, academic detailing, audit, and feedback.

That physicians do not always counsel patients to stop smoking, do not always order screening mammograms when women need them, do not prescribe angiotensin-converting enzyme inhibitors for all patients with congestive heart failure — despite guidelines from multiple organizations recommending these practices — emphasizes the barriers that physicians face in changing behavior.²⁵ Beyond *knowing* what to do, clinicians also encounter *attitudinal* barriers in accepting the validity of recommendations, *implementation* barriers in operationalizing recommendations in practice, and *reinforcement* barriers in maintaining a commitment to the intervention over time.²⁶ That a guideline recommends a service matters little if the clinician disputes it, cannot provide it (due to lack of time, equipment, or support from the health care system), or forgets that it is indicated.

The new frontier in EBM lies in developing effective strategies for translating evidence into practice. The most promising approaches tailor implementation strategies to the type of barriers involved, eg, local opinion leaders and academic detailing when attitudes are the obstacle, changing office operations, redesigning order forms, skill building, decision support tools to enhance implementation, and computerized reminder systems and

other prompts to provide reinforcement.²⁷ These measures, when used to apply the growing wealth of data from clinical research, are most likely to fulfil the promise of EBM to improve the efficacy and equity of health care.

References

1. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med.* 1997;126:376-380.
2. L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med.* 1987;107:224-233.
3. Bero L, Rennie D. The Cochrane Collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. *JAMA.* 1995;274:1935-1938.
4. Chassin MR, Koseoff J, Park RE, et al. Does inappropriate use explain geographic variations in the use of health care services? A study of three procedures. *JAMA.* 1987;258:2533-2537.
5. Institute of Medicine. *To Err is Human: Building a Safer Health System.* Washington, DC: National Academy Press; 1999.
6. Sackett DL, Rosenberg WM, Gray JA, et al. Evidence based medicine: what it is and what it isn't. *BMJ.* 1996;312:71-72.
7. Feinstein AR, Horwitz RI. Problems in the "evidence" of "evidence-based medicine." *Am J Med.* 1997;103:529-535.
8. Tanenbaum SJ. "Medical effectiveness" in Canadian and US health policy: the comparative politics of inferential ambiguity. *Health Serv Res.* 1996;31:517-532.
9. Kane RL. Creating practice guidelines: the dangers of over-reliance on expert judgment. *J Law Med Ethics.* 1995;23:62-64.
10. Eddy DM. Variations in physician practice: the role of uncertainty. *Health Aff (Millwood).* 1984;3:74-89.
11. The Cardiac Arrhythmia Suppression Trial (CAST) Investigators. Preliminary report: effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction. *N Engl J Med.* 1989;321:406-412.
12. Mullan F, Jacoby I. The town meeting for technology: the maturation of consensus conferences. *JAMA.* 1985;254:1068-1072.
13. Brook RH, Chassin MR, Fink A, et al. A method for the detailed assessment of the

appropriateness of medical technologies: a RAND Note. N-3376-HHS. Santa Monica, Calif: RAND; 1991.

14. Cook D, Giacomini M. The trials and tribulations of clinical practice guidelines. *JAMA*. 1999;281:1950-1951.

15. Woolf SH. An organized analytic framework for practice guideline development: using the analytic logic as a guide for reviewing evidence, developing recommendations, and explaining the results. In: McCormick KA, Moore SR, Siegel RA, eds. *Clinical Practice Guideline Development: Methodology Perspectives*. AHCPR Pub No 95-0009. Rockville, Md: Agency for Health Care Policy and Research; 1995:105-113.

16. Cook DJ, Guyatt GH, Laupacis A, et al. Clinical recommendations using levels of evidence for antithrombotic agents. *Chest*. 1995;108(4 suppl):227S-230S.

17. Elwood JM. *Critical Appraisal of Epidemiological Studies and Clinical Trials*. 2nd ed. Oxford, NY: Oxford University Press; 1998.

18. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA*. 1993;270:2598-2601.

19. Moher D, Jadad AR, Tugwell P. Assessing the quality of randomized controlled trials: current issues and future directions. *Int J Technol Assess Health Care*. 1996;12:125-208.

20. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*. 1996;17:1-12.

21. George JN, Woolf SH, Raskob GE. The evidence-based analysis of treatment for chronic myeloid leukemia: an introduction to its methods and clinical implications. *Blood*. 1999;94:1515-1516.

22. Woolf SH. Shared decision-making: the case for letting patients decide which choice is best. *J Fam Pract*. 1997;45:205-208.

23. Woolf SH. Do clinical practice guidelines define good medical care? The need for good science and the disclosure of uncertainty when defining "best practices." *Chest*. 1998;113:166S-171S.

24. Davis DA, Thomson MA, Oxman AD, et al. Evidence for the effectiveness of CME: a review of 50 randomized controlled trials. *JAMA*. 1992;268:1111-1117.

25. Grol R. Personal paper. Beliefs and evidence in changing clinical practice. *BMJ*. 1997;315:418-421.

26. Woolf SH, Grol R, Hutchinson A, et al. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines.

BMJ. 1999;318:527-530.

27. Bero LA, Grilli R, Grimshaw JM, et al. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. *BMJ*. 1998; 317: 465-468.