Evidence-based care: 2. Setting guidelines: How should we manage this problem?

Evidence-Based Care Resource Group

There are four steps in determining how to manage a clinical problem. The first is to formulate questions that are answerable; the second is to locate and synthesize the evidence needed to answer the questions; the third is to estimate the expected benefits, harms and costs of each option based on the evidence; and the fourth is to judge the relative value of the expected outcomes to conclude whether the benefits are worth the harms and costs. It is impractical to repeat these steps for every clinical decision. Therefore, implicitly or explicitly, physicians rely on guidelines, "rules" that simplify decision making about complex problems. If the methods used to develop a guideline are not explicit it is difficult or impossible to know how much confidence to place in it. Therefore, for common and important clinical problems, physicians should rely on guidelines that are systematically developed using explicit methods.

Il y a quatre étapes à suivre pour décider comment traiter un problème clinique. Il faut d'abord formuler des questions auxquelles il est possible de répondre, ensuite, trouver et résumer les données nécessaires à cette fin, troisièmement, évaluer les avantages, les préjudices et les coûts prévus de chaque solution possible en fonction des données et, enfin, établir la valeur relative des résultats attendus afin de déterminer si les avantages valent les préjudices et les coûts en jeu. Ces étapes ne sont pas pratiques dans le cas de chaque décision clinique à prendre. C'est pourquoi les médecins comptent implicitement ou explicitement sur des lignes directrices, des «règles» qui simplifient la prise de décisions au sujet de problèmes complexes. Si les méthodes d'élaboration d'une ligne directrice ne sont pas claires, il est difficile ou impossible de savoir dans quelle mesure il faut s'y fier. C'est pourquoi, dans le cas des problèmes cliniques communs et importants, les médecins devraient compter sur des lignes directrices élaborées de façon systématique à l'aide de méthodes explicites.

Should you treat a 75-year-old woman with a systolic blood pressure of 180 mm Hg? Should you refer a 65-year-old man with symptomatic, benign hypertrophy of the prostate for surgery? What should you tell a woman of 55 who wants to know whether she should start hormone replacement therapy? How should you manage a case of acute myocardial infarction in a 45-year-old man? Should you order a glucose tolerance test for a 35-year-old woman who is in her 26th week of pregnancy? Should you refer a 25-year-old man with

Members: Drs. Andrew D. Oxman (chair and principal coauthor), David A. Davis, John W. Feightner (principal coauthor), Neil V. Finnie and Brian G. Hutchison, Ms. Sandy Lusk and Drs. Peter J. MacDonald, Ron G. McAuley and John W. Sellors

Drs. Oxman, Feightner, Hutchison, McAuley and Sellors are in the departments of Family Medicine and of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ont. Drs. Davis, Finnie and MacDonald and Ms. Lusk are in the Department of Family Medicine, McMaster University. Dr. Oxman was the recipient of a Career Scientist Award from the Ontario Ministry of Health.

Reprint requests to: Evidence-Based Care Resource Group, Department of Family Medicine, Rm. 2V6, Health Sciences Centre, McMaster University, 1200 Main St. W, Hamilton, ON L8N 3Z5; fax (905) 528-5337

This article is the second in a series of five that began in the Apr. 15, 1994, issue of CMAJ.

← For prescribing information see page 1515

acute low-back pain for spinal manipulation? What should you recommend to a sexually active girl of 15? How should you manage acute otitis media in a child of 5?

Clinicians are confronted daily with decisions such as these. They must decide which questions to ask patients, what to include in a physical examination, which diagnostic tests to order, what to tell patients, which interventions to recommend or use and what follow-up is needed. To aid in making these decisions physicians, implicitly or explicitly, must rely on guidelines — simple decision rules for resolving complex problems.[1,2]

For example, a decision on whether to treat systolic hypertension in elderly patients entails four key steps (Fig. 1). First, an answerable question must be posed. Second, the evidence needed to answer the question must be located and critically reviewed. Third, the expected benefits (e.g., reduced risk of stroke), harms (e.g., side effects of drugs) and costs of treatment must be estimated. Finally, a judgement about the relative value of the expected benefits, harms and costs must be made. If the treatment of 100 patients for 5 years prevents 3 of them having a stroke, 1 having a heart attack and 2 or 3 suffering congestive heart failure yet will result in 3 or 4 patients having intolerable side effects,[3] is systolic blood pressure in elderly patients worth treating?

It is impractical and unreasonable to analyse clinical decisions, especially common ones, repeatedly this way. Instead, physicians rely on guidelines such as: "Hypertension therapy should be prescribed for patients 60 years of age and older with isolated systolic hypertension (diastolic blood pressure less than 90 mm Hg and

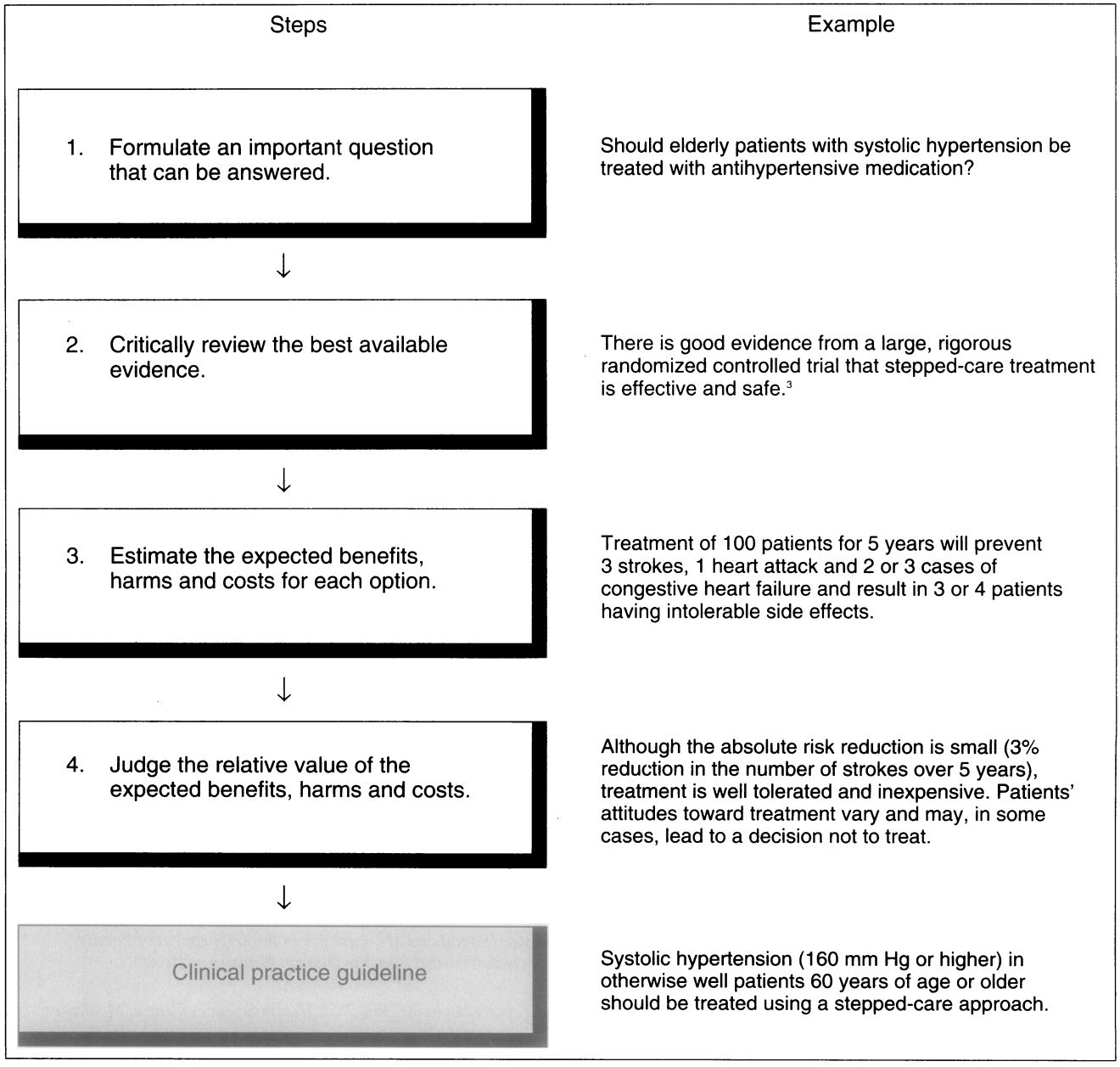

Fig. 1: Key steps in developing a clinical practice guideline.

systolic blood pressure 160 mm Hg or higher)."[4] With the use of such a guideline a complex problem becomes manageable.

Of course, the quality of the care depends on the quality of the guidelines used. In this article we will outline some important considerations in each of the four key steps in determining appropriate guidelines for clinical practice.

These considerations also apply to guidelines developed with an implicit approach or nonsystematically;[5] however, the considerations shown in Fig. 1 should be addressed systematically to avoid potential mistakes that occur in nonsystematically derived guidelines,[6,7] and methods for setting guidelines should be explicit to enable critical appraisal of the guidelines' validity.[8]

What is the problem?

After deciding that a problem is important[9] physicians should characterize the patients to whom the decisions apply and clarify the options and outcomes of interest. In the first article of this series we identified two types of problems: how to manage a clinical condition and how to improve the delivery of health services. In both cases a decision must be made. There are always at least two options (e.g., to use or not to use an intervention) and often several. There is at least one outcome of interest, and typically there are several, including potential benefits, harms and costs. Examples of the target patients, options and outcomes for various clinical problems are shown in Table 1. The next three steps become clearer after the problem has been specified in this way.

What is the evidence?

Occasionally the benefits of an intervention are so clear, and the harms and costs so small, that there is little or no need for rigorous evaluation (e.g., transfusion for massive blood loss, injection of epinephrine for anaphylaxis, administration of penicillin for pneumococcal pneumonia and reduction of a dislocated elbow in a toddler). The effectiveness of most care is not so obvious, and rigorous evaluations are needed to determine whether the perceived benefits are real and worth while.

To ensure that good research is translated into good clinical decisions clinicians must be informed consumers.[10] A large amount of medical information is not supported by valid research, including some articles published in prestigious medical journals and recommendations made by leading authorities.[11,12] Clinicians must be selective about what they read and heed to ensure that it is applicable and valid. The opinion of authorities and one's own clinical experience are not adequate to validate the results of research.[12–14] To make informed decisions clinicians must be informed users of medical research and have the ability to appraise it critically.[15]

This does not mean that each clinician must review all of the original research relevant to his or her practice. For much, if not most, of what clinicians do they have to rely on others to locate, critically appraise and synthesize the research in the form of a systematic review[16,17] or practice guidelines.[7]

Table 1: Examples of target patients, options and outcomes for various clinical problems

Problem	Target patients	Option	Outcomes
Therapy	Patients with acute asthma	Oral corticosteroid therapy	Benefit: Improved functional status Harm: Side effects Costs: Cost of drugs, savings from reduced hospital use
Prevention	Patients who have had a myocardial infarction	ASA* prophylaxis	Benefit: Reduced risk of cardiovascular (CV) events Harm: Side effects Costs: Costs of drugs and side effects, savings from CV events prevented
Diagnosis	Adult patients with sore throat	Throat swab of mucous sample for culture	Benefits: Increase in indicated treatment with penicillin, reduction in unnecessary treatment and side effects Harm: Delay in starting indicated treatment and relieving symptoms Costs: Costs of culture and treatment
Screening	Asymptomatic adult patients	Screening for total serum cholesterol level	Benefit: Reduced risk of CV events Harms: Adverse consequences of follow-up, labelling Costs: Cost of test and follow-up, savings from CV events prevented
Health services delivery	Patients seen in primary care practice	Computerized reminders for periodic health examinations	Benefit: Improved delivery of effective preventive services Harm: Deterioration in delivery of services for which reminders are not provided Costs: Cost of computerization, time spent contacting patients and performing examinations

*ASA = acetylsalicylic acid.

Criteria for screening articles about therapy, prevention, diagnostic tests and health service studies, review articles and clinical practice guidelines are summarized in Table 2. If an article has met these criteria it is more likely to be valid.

In some areas, such as general internal medicine,[19] obstetrics[20] and neonatology,[21] there are information sources that have already applied similar criteria. The *ACP Journal Club* prepares structured summaries of articles on general internal medicine selected from more than 40 journals that meet validity criteria like those in Table 2.[19] Similar efforts have been started or are planned for other disciplines (Dr. R. Brian Haynes, editor, *ACP Journal Club*, and in the departments of Clinical Epidemiology and Biostatistics and of Medicine, McMaster University, Hamilton, Ont.: personal communication, 1994), including pediatrics.[22] The Cochrane Pregnancy and Childbirth Database consists of regularly updated systematic reviews of controlled trials of obstetric care.[23] The Cochrane Collaboration was formed recently to meet the need for systematic, up-to-date reviews of controlled trials of all forms of health care and to make this information readily available to clinicians and other decision makers at all levels of health care systems.[24]

Although these resources make the tasks of coping with the medical literature and practising evidence-based medicine more manageable, they do not eliminate the need for clinicians to be critical consumers of scientific information. However, we have found that, with a little experience, critical reading not only provides the basis for improved quality of care but also is empowering and fun.[10]

Table 2: Criteria for identifying valid articles*

Type of article	Criteria
Trial of therapy or preventive intervention	Random assignment of patients to treatment groups; accountability of all patients at end of trial
Trial of diagnostic test	Independent, blind comparison with a gold standard; appropriate number of patients included
Trial of health service intervention	Random assignment of patients to comparison groups; reporting of clinically important outcomes for all participants
Review	Clearly focused question addressed; appropriate criteria used to select articles for inclusion
Clinical practice guidelines	Options and outcomes clearly specified; explicit process to identify, select and combine evidence

*Adapted from Oxman, Sackett and Guyatt.[15] Guides for critical appraisal and application of the medical literature have been described in detail elsewhere.[15,18,19]

What are the expected outcomes?

After valid research is located or the quality of the available evidence is determined the next step is to estimate the expected consequences of the options being considered. In general there are three categories of outcomes: expected benefits, potential harms and costs.

For therapeutic and preventive interventions it is useful to consider the effects of treatment in absolute terms. For instance, the clinical importance of a reduction in relative risk depends on the baseline risk and the severity of the consequences of lack of treatment. If the baseline risk of a stroke in a certain population is 50%, a 20% reduction in relative risk would result in the prevention of one stroke for every 10 patients treated and would probably be considered worth while. On the other hand, if the baseline risk is 0.1%, a 20% reduction in risk would result in the prevention of one stroke for every 5000 patients. Results of research are often reported as relative risks or similar measures, which do not convey clinical importance.[25] However, it is possible to translate results into more clinically relevant measures, such as the number needed to treat, by considering the baseline risk.[26]

For diagnostic tests the cardinal question is whether the test results can affect the estimated probability that a patient has a condition sufficiently to influence clinical action. If this probability is very high or very low, ordering a test is generally less likely to add useful information. Just as it is important to consider the baseline risk and the severity of consequences for therapeutic and preventive interventions, it is important to consider the before-test probability, the value of subsequent clinical actions, the potential harms and benefits, and the costs for a diagnostic intervention (Fig. 2). For example, if adverse consequences are associated with false-positive results of screening tests (e.g., screening for occult blood for early detection of colorectal cancer) the test must have good specificity. Conversely, if only minor adverse consequences are associated with false-positive results of screening tests (e.g., screening for phenylketonuria) but a missed diagnosis is associated with severe consequences, the test's sensitivity is more important than its specificity.

When considering costs, societal costs and costs borne by the patient must be distinguished. In general, costs borne by the patient should be considered in relation to his or her situation. However, for a clinician to make a judgement about societal costs within the context of a consultation is in conflict with his or her role in providing care. For example, a clinical practice guideline may recommend against routine ultrasound screening in pregnancy because the proven benefits are small and the costs are substantial.[27] However, such a guideline should be determined in relation to a group or population of patients. In the absence of such guidelines, clinicians are in

conflict if they try to consider societal costs when making decisions regarding specific patients.

What are the trade-offs?

The last step is to weigh the benefits against the harms and costs. This typically involves comparing widely different outcomes, such as reduced risk of illness, side effects of drugs, the patient burden of taking medication daily and the cost of drugs. Clearly, the value attached to a major stroke is different from that assigned to a mild side effect such as dizziness. Athough it is obvious that these outcomes have different values, their relative values are less apparent. They are also likely to vary from one patient to another. For example, women with breast cancer provided with the same information about the benefits and harms of chemotherapy have different preferences and make different decisions about their treatment.[28] The same is true for men considering surgical treatment of symptomatic, benign hypertrophy of the prostate[29] and women considering hormone replacement therapy.[30]

In general, decisions that depend on the relative value attached to the main consequences of a decision should be left to the patient. In such situations the clinician's role is to assist the patient in arriving at a decision, not to make a decision for the patient.

However, sometimes patients want physicians to make decisions for them.[31] In these and other circumstances societal preferences can be important guides for clinical decisions, particularly when the benefits are small relative to societal costs. Such preferences are also important in clinical decisions about problems that affect people other than the patient (e.g., some communicable diseases and psychiatric problems) or that affect children, about terminally ill patients and about unsolicited preventive interventions.

For many clinical decisions it is impractical and unnecessary to quantify the values attached to the consequences. However, in the same way that it is important for evidence and expected outcomes to be explicit, judgements of preference in weighing benefits against harms and costs should also be explicit.[8,32]

How should the problem be managed?

Although clinical decisions often seem black and white (e.g., to prescribe a drug or not) there is usually a range of options from always doing something to never doing it (Fig. 3). In addition, there are almost always caveats when implementing a guideline. For example, contraindications of or exemptions from preventive and therapeutic interventions are common. Sometimes a

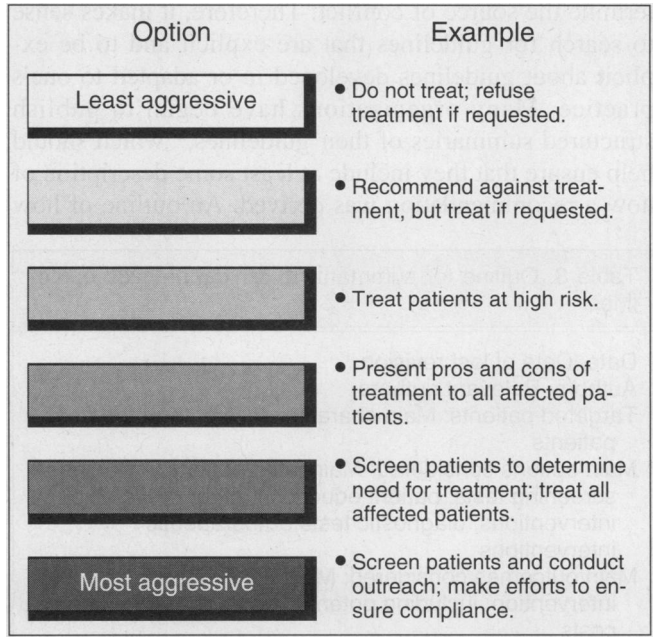

Fig. 3: "Shades of grey" in treatment recommendations contained in clinical practice guidelines.

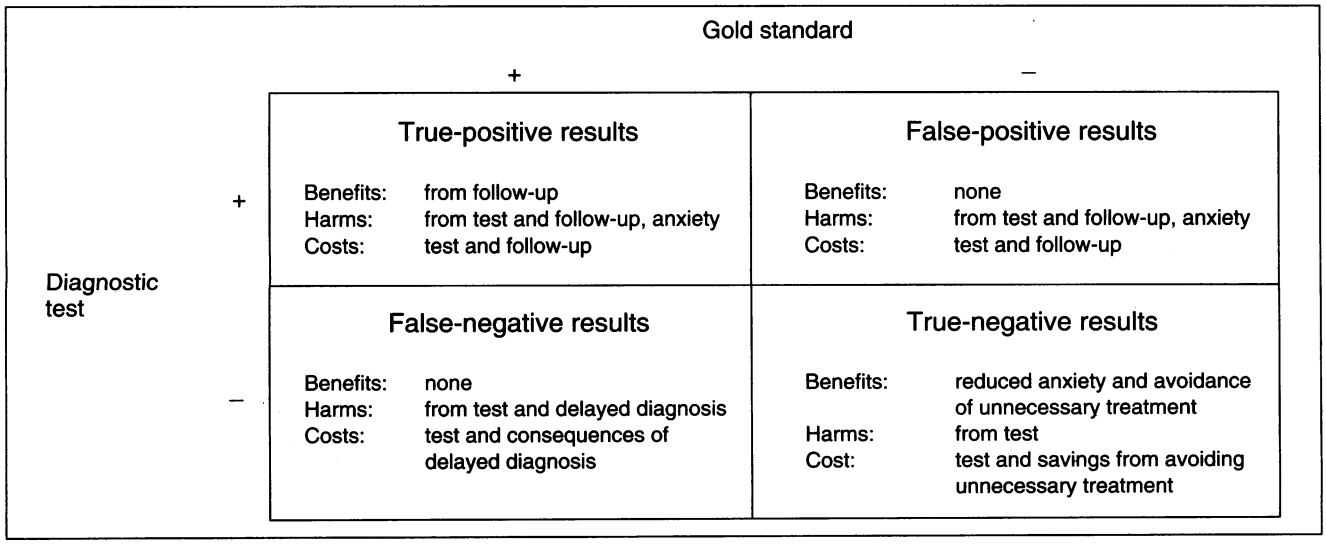

Fig. 2: Benefits, harms and costs of diagnostic tests.

guideline must be adjusted to take into account patient characteristics (e.g., socioeconomic status) or characteristics of the local setting (e.g., availability of resources).

If clinicians focus on common problems, as we have suggested,[9] they are likely to find more than one guideline for interventions. Whether clinicians develop their own guideline or determine whether a guideline is valid and applicable in their practice, they must recognize any conflicts with other guidelines for the same clinical problem. If possible, clinicians should identify the extent to which differences in recommendations are due to differences in how the problem was framed, how the evidence was assessed, how outcomes were estimated or how judgements about preferences were made.

If a clinical practice guideline is not explicit about each of these steps it is difficult, if not impossible, to determine the source of conflict. Therefore, it makes sense to search for guidelines that are explicit and to be explicit about guidelines developed in or adapted to one's practice. Many organizations have begun to publish structured summaries of their guidelines,[33] which should help ensure that they include at least some description of how a recommendation was derived. An outline of how we summarize clinical practice guidelines is shown in Table 3.

Table 3: Outline for summarizing clinical practice guidelines

Date: Date of last revision
Authors: Relevant authors
Targeted patients: Main characteristics of targeted patients
Main options considered: Main interventions, including screening tests, patient education, other preventive interventions, diagnostic tests or therapeutic interventions
Main outcomes considered: Main consequences of intervention, including potential benefits, harms and costs
Evidence: Brief summary of main evidence including only key citations (best overviews and critical studies) and indication of strength of the evidence
Recommendation: Succinct statement (one to three sentences) of proposed policy
Expected benefits, harms and costs: Quantitative or qualitative estimate of main consequences that led to recommendation
Preference judgements: Identification of key considerations concerning patient preferences for the expected outcomes
Other guidelines: Reference to other guidelines and consistency with them
Dissenting opinions: Identification of any disagreement among professional staff and main source of disagreement
Caveats: Identification of any important caveats
Implementation: Specification of primary implementation strategies
Evaluation: Statement of any plans for evaluating impact of clinical policy
Information sources: Main strategies used to identify evidence
Key references: Applicable references

Because medical knowledge and practice environments evolve continually, guidelines have a "shelf life" after which they should be reassessed. More important, the determination of interventions for a significant problem is of little benefit to our patients if guidelines just sit on a shelf. In the next two articles in this series we will discuss strategies for measuring what physicians are doing and implementing guidelines to close any gaps between what they should be doing and what they are doing.

The next article will focus on decisions made in measuring clinical performance: what to measure, whether the information needed to make the measurement is available, how to select an appropriate sample of patients, how to collect the information needed and how to interpret the information collected.

References

1. Battista RN, Hodge MJ: Clinical practice guidelines: between science and art. *Can Med Assoc J* 1993; 148: 385–389
2. Oxman AD: Coordination of guidelines development. *Can Med Assoc J* 1993; 148: 1285–1288
3. SHEP Cooperative Research Group: Prevention of stroke by antihypertensive drug treatment in older persons with isolated systolic hypertension: final results of the systolic hypertension in the elderly program (SHEP). *JAMA* 1991; 265: 3255–3264
4. Haynes RB, Lacourcière Y, Rabkin SW et al: Report of the Canadian Hypertension Society Consensus Conference: 2. Diagnosis of hypertension in adults. *Can Med Assoc J* 1993; 149: 409–418
5. Eddy DM: Clinical policies and the quality of clinical practice. *N Engl J Med* 1982; 307: 343–347
6. Eddy DM: Practice policies: guidelines for methods. *JAMA* 1990; 263: 1839–1841
7. Hayward RSA, Laupacis A: Initiating, conducting and maintaining guidelines development programs. *Can Med Assoc J* 1993; 148: 507–512
8. Eddy DM: Guidelines for policy statements: the explicit approach. *JAMA* 1990; 263: 2239–2243
9. Evidence-Based Care Resource Group: Evidence-based care: 1. Setting priorities: How important is this problem? *Can Med Assoc J* 1994; 150: 1249–1254
10. Evidence-Based Medicine Working Group: Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA* 1992; 268: 2418–2425
11. Williamson JW, Goldschmidt PG, Colton T: The quality of medical literature: an analysis of validation assessments. In Bailar JC, Mosteller F (eds): *Medical Uses of Statistics*, NEJM Books, Waltham, Mass, 1986: 370–391
12. Antman EM, Lau J, Kupelnick B et al: 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
13. W illiamson JW, German PS, Weiss R et al: Health science information management and continuing education of physicians. *Ann Intern Med* 1989; 110: 151–160
14. Oxman AD, Guyatt GH: The science of reviewing research. *Ann N Y Acad Sci* 1993; 703: 125–134
15. Oxman AD, Sackett DL, Guyatt GH for the Evidence-Based Medicine Working Group: Users' guides to the medical literature. I: How to get started. *JAMA* 1993; 270: 2093–2095

16. Oxman AD, Guyatt GH: Guidelines for reading literature reviews. *Can Med Assoc J* 1988; 138: 697–703
17. Oxman AD, Cook DJ, Guyatt GH for the Evidence-Based Medicine Working Group: Users' guides to the medical literature. VI: How to use an overview. *JAMA* (in press)
18. Sackett DL, Haynes RB, Guyatt GH et al: The interpretation of diagnostic data. In *Clinical Epidemiology: a Basic Science for Clinical Medicine*, Little, Brown and Company, Boston, 1991: 69–152
19. Purpose and procedure. *ACP J Club* 1991; Jan–Feb 1991: A6–A7 (*Ann Intern Med* 114, suppl 1)
20. Chalmers I, Hetherington J, Elbourne D et al: Materials and methods used in synthesizing evidence to evaluate the effects of care during pregnancy and childbirth. In Chalmers I, Enkin M, Keirse MJNC (eds): *Pregnancy and Childbirth*, Oxford University Press, Oxford, England, 1989: 39–66
21. Sinclair JC, Bracken MB (eds): *Effective Care of the Newborn Infant*, Oxford University Press, Oxford, England, 1992: 3–18
22. Moyer V: Spreading the Journal Club net — pediatrics. *Clin Epidemiol News* 1993; 12: 13
23. Enkin MW, Keirse MJNC, Renfrew MJ et al (eds): *Cochrane Pregnancy and Childbirth Database [derived from the Cochrane Database of Systematic Reviews]*, Cochrane Updates on Disk, Update Software, Oxford, England, 1993: disk issue 2
24. *The Cochrane Collaboration* [brochure], UK Cochrane Centre, National Health Service Research and Development Programme, Oxford, England, 1993
25. Naylor CD, Chen E, Strauss B: Measured enthusiasm: Does the method of reporting trial results alter perceptions of therapeutic effectiveness? *Ann Intern Med* 1992; 117: 916–921
26. Laupacis A, Sackett DL, Roberts RS: An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988; 318: 1728–1733
27. Neilson JP: Routine ultrasonography in early pregnancy. In Enkin MW, Keirse MJNC, Renfrew MJ et al (eds): *Cochrane Pregnancy and Childbirth Database [derived from the Cochrane Database of Systematic Reviews]* (review no 03872, Mar 24, 1993), Cochrane Updates on Disk, Update Software, Oxford, England, 1993: disk issue 2
28. Levine MN, Gafni A, Markham B et al: A bedside decision instrument to elicit a patient's preference concerning adjuvant chemotherapy for breast cancer. *Ann Intern Med* 1992; 117: 53–58
29. Barry MJ, Mulley AG, Fowler FJ et al: Watchful waiting vs immediate transurethral resection for symptomatic prostatism: the importance of patients' preferences. *JAMA* 1988; 259: 3010–3017
30. American College of Physicians: Guidelines for counseling postmenopausal women about preventive hormone therapy. *Ann Intern Med* 1992; 117: 1038–1041
31. Degner LF, Sloan JA: Decision making during serious illness: What role do patients really want to play? *J Clin Epidemiol* 1992; 45: 941–950
32. Eddy DM: Anatomy of a decision. *JAMA* 1990; 263: 441–443
33. Hayward RSA, Tunis SR, Wilson MC et al: More informative abstracts of articles describing clinical practice guidelines. *Ann Intern Med* 1993; 118: 731–737