Standards, guidelines and clinical policies

Health Services Research Group

The explosion in the amount of medical information available has made it increasingly difficult for physicians to remain abreast of important discoveries, let alone incorporate them into routine clinical practice. Furthermore, studies reporting significant and well-validated advances are scattered throughout an enormous literature of uneven quality.[1] One goal of establishing clinical policies is to provide practitioners with a structured, evidence-based approach to the provision of procedures and services.

The assumption that large variations in clinical practice patterns are indicative of parallel variations in the quality of medical care[2] may not be well founded, but some proportion of unnecessary services has been identified in many areas of practice.

Providers, citizens and governments alike are concerned with cost containment; although this should not be the primary goal of policy development, avoidance of unnecessary care may enhance efficiency.

Methods to implement and evaluate clinical practice policies lag behind the enthusiasm for setting them, and the obstacles that confound their adoption go unidentified or unsurmounted. Policies vary with the clinical domain, the quality of the relevant evidence, the level of agreement among experts and the goals of those setting the policies. We review the types of clinical policies and their influence on clinical practice, and we discuss future directions in this field.

Standards, practice guidelines and policies

Standards of quality are "authoritative statements of 1) minimum levels of acceptable performance or results or 2) excellent levels of performance or results, or 3) the range of acceptable performance."[3] Practice guidelines are defined as "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances."[3] They direct decisions made in particular circumstances but may be set aside by individual considerations and choices. In contrast, minimum standards of care define boundaries that distinguish acceptable from unacceptable practice. Whereas standards may be set forth for all aspects of structure, process and outcome, practice guidelines tend to pertain primarily to the processes of health care. None the less, there is disagreement as to the division between guidelines and standards.

A policy is "a definite course or method of action selected from among alternatives and in light of given conditions to guide and determine present and future decisions."[4] We will use "policy" to refer either to guidelines or to standards.

Donabedian[5] has discussed issues of structure, process and outcome in ensuring high-quality health care. Similarly, clinical policies may address structures (e.g., well-maintained crash carts on hospital wards and the need for a physician to meet training requirements before receiving admitting privileges), processes (e.g., the use of mammography for screening purposes and indications for thrombolytic drugs in acute myocardial infarction) or outcomes (e.g., maximum acceptable complication rates for maintenance of surgical privileges).

Guidelines and standards may be promulgated in various guises. Algorithmic approaches prescribe

Members: Drs. Antoni S.H. Basinski (principal coauthor), Department of Family and Community Medicine; C. David Naylor (principal coauthor), Department of Medicine; Marsha M. Cohen, Department of Health Administration; Lorraine E. Ferris, Department of Behavioural Science; J. Ivan Williams, Department of Preventive Medicine and Biostatistics; and Hilary A. Llewellyn-Thomas, Faculty of Nursing, University of Toronto, Toronto, Ont.

The Health Services Research Group is part of the Clinical Epidemiology Unit, Sunnybrook Health Science Centre, Toronto, Ont.

Reprint requests to: Health Services Research Group, Clinical Epidemiology Unit, Rm. A443, Sunnybrook Health Science Centre, 2075 Bayview Ave., Toronto, ON M4N 3M5

a collection of decision paths encompassing diagnostic or treatment options, or both (e.g., the cholesterol guidelines of the US National Cholesterol Education Program.[6]) Alternatively, standards may consist of a number of necessary and sufficient conditions delineating the conduct of good medical care.

Appropriateness, necessity and urgency

A medical procedure or service may be described as appropriate if its benefits exceed its risks sufficiently to make it worth performing.[7] The Rand approach has used panels of expert clinicians to rate hypothetic clinical situations as "appropriate," "equivocal" (now termed "uncertain") or "inappropriate" for a given procedure.[7-9] Hypothetic situations are then matched with actual cases. Assessments are made retrospectively for research purposes in chart audits and prospectively for utilization management by third-party payers in the United States.

Most assessments of appropriateness have focused on overuse,[7] but some studies imply that there are problems with underuse[10] or limitations in access to care. Appropriateness ratings are tantamount to guidelines, in that a procedure may in some circumstances be appropriate but not absolutely necessary. Rand researchers are now working to develop "necessity" ratings that define a basic standard below which underuse and impaired access are present.

Given the ubiquity of waiting lists for procedures it may be helpful to determine also the urgency of a procedure. For example, recent efforts have been undertaken to define the urgency of the need for coronary artery surgery.[11] The relation between appropriateness, necessity and urgency is still being explored.

Needs and goals for policies

Ideally, policies are set in response to a perceived need. However, many policies are developed without clear delineation of intent and goals, and this may lead to misunderstandings between policy sponsors and target groups of clinicians.

When goals are set they must be realistic and be consistently addressed by the guidelines.[12] A distinction can be drawn between efficacy (the estimated benefit, assuming that the recommended practice is universally applicable and that physicians and patients comply with it) and effectiveness (the benefit if the practice has reduced applicability or there is reduced compliance). Thus, estimates of the potential benefit of the universal detection and management of hypercholesterolemia were of a 10% decrease in the rates of death from coronary heart disease. Estimates based on empiric evidence of reduced follow-up and compliance were of a 2% decrease in the death rate.[13] Guidelines that use optimistic extrapolations will disappoint sponsors and frustrate patients and practitioners alike.

Setting policies

The debate surrounding practice policies has focused on the issue of who sets policies and how.[14] To date, practice policies have largely emanated from interest groups, specialist bodies, practice organizations and "independent agencies." The conflict over who sets the policies in part reflects tensions between professional autonomy and the prerogative of consumer and third-party groups to representation. Specialist bodies and interest groups may have a vested interest in promoting particular procedures or approaches, whereas insurers, governments and nonpractitioners may seek to limit the use of the same procedures and services.

Many have called for the establishment of an explicit process for setting policies that are based on scientific evidence.[15-21] With clear links between the evidence and the recommendations, distinctions are brought to light between what is known and what is assumed and between expert opinion and scientific knowledge.[16]

A number of approaches have been developed for setting explicit policies. The Canadian Task Force on the Periodic Health Examination[22] and more recently the US Preventive Health Task Force[23] have based their recommendations on a formalized grading of the evidence and the strength of recommendations for or against a clinical practice. The Rand approach is to provide systematic reviews of the literature to a panel of experts, who then engage in a modified Delphi process to rate the appropriateness of procedures or services.

Even these excellent models founder because of limited evidence. Many procedures have not been and never will be tested in randomized clinical trials.[21] When evidence from such trials exists it is often incomplete, and generalizations cannot be made. Thus, it is necessary to weigh and synthesize evidence of various types in setting most clinical policies. Value judgements and inference are inevitable. In their absence, frameworks requiring faithful adherence to a methodologic catechism of policy development (such as that proposed by the Canadian Task Force on the Periodic Health Examination) may become a procrustean bed on which common sense is racked in the name of methodologic purity. On the other hand, the Rand approach formalizes the degree of agreement (or disagreement) in the opinion of a panel of experts, and the links to the evidence may become tenuous indeed.

Consensus and its paradox

Consensus is defined as the "agreement in opinion, the collective unanimous opinion of a number of persons."[24] Most policy-setting groups aim at providing statements based on the proclamations of consensus groups and expert panels. The impetus for such reports lies in a quest for assurance that "good" is being done in an environment of complexity and uncertainty.

There is no consensus on how consensus is achieved; few formal mechanisms exist for dealing with disagreement.[25] Yet genuine unanimity is only feasible when the implications and the weight of the evidence are uncontested — a rarity in clinical medicine. In such circumstances, consensus statements are important primarily in education and standard-setting; that is, they legitimize the "collective wisdom" and thereby help close the gap between indisputable evidence and the incorporation of that evidence into practice. When the evidence is incomplete there will be legitimate variation in the inferences that are drawn from it. Common ground in competing interpretations may serve to define a basic standard of care; beyond that, practitioners must be free to choose among competing guidelines without coercion. Last, to blunt the antiscientific features of consensus statements it may be best to insist on clear information about what is known and what is not. The implications of both knowledge and uncertainty ought to guide the practitioner.

Relation to quality improvement

Formal efforts to improve medical care require the establishment of standards for clinical practice against which current practice can be measured. Standards may relate not only to when but also to how a service is performed. They may range from minimal standards for specific services to standards of excellence over broad domains to provide targets for improved quality of care. The model of continual quality improvement of health care is predicated on assessment and accountability.[26,27] Assessment calls for an organization to set standards for structure, process and outcome and for practice patterns to be measured against the defined standard. Accountability ensures that action is taken to respond to practices that fall short of defined standards.

Dissemination and implementation

Currently guidelines are released and promoted in peer-reviewed journals, mailings, press releases and presentations at major meetings. Compliance with guidelines may be promoted by training sessions and educational materials.[28] However, the strongest predictor of congruent practitioner behaviour after the dissemination of practice guidelines is prior practice behaviour compatible with the recommendations.[29,30] Hence, successful guidelines may codify existing practices rather than change attitudes.

If existing practices are less in accord with guidelines simple dissemination may not suffice to change them. In the United States, therefore, guidelines are commonly linked to other factors (for instance, quality assurance manoeuvres or reimbursement) to effect change in provider behaviour.[31] Preadmission authorization for surgery is required by many US insurers, and payment is withheld if standards are not met.[32] Formal standards may also be made mandatory by regulation or law. In Canada, standards of practice within a quality assurance program are required of clinics and day hospitals under the Independent Health Facilities Act of Ontario.[33]

Factors influencing the acceptability of a policy include the representation of the user group on the setting body and the perceived "ownership" of the distilled opinion. Although the authoritativeness of the sponsoring agency or source has been claimed to be a strong predictor of the influence of a given guideline on practice patterns,[30] locally developed guidelines may also be very influential because of increased acceptance and peer pressure. For example, locally produced guidelines for a reduction in the testing of hospital patients were generally accepted and led to a 20% to 42% reduction in the use of laboratory tests and in radiologic and electrocardiographic examinations.[34]

In Ontario the effect of a widely distributed and nationally endorsed consensus statement on the use of cesarean section was assessed in a survey of obstetricians and by estimation of actual practice. Most obstetricians agreed with the guidelines. In fact, attitudes toward the use of cesarean section were congruent with the recommendations even before their release. However, knowledge of the content of the recommendations was poor. Independent data revealed that the rates of cesarean section were substantially higher than the rates estimated by obstetricians, and they showed only a slight change from the previous upward trend.[35]

In contrast, a quality improvement initiative to reduce the number of cesarean deliveries in an inner-city hospital was successful.[36] Participation by physicians was voluntary. The program included a second opinion, objective criteria for the most common indications for cesarean section and a detailed review of all cesarean sections and of individual physicians' rates of performing them. During the first 2 years of the program the rate of cesarean sections decreased from 17.5% to 11.5% of deliveries

without adverse effects on mothers or their infants.[36]

There are a number of factors that may contribute to the differences observed in these studies. In the first case there were no incentives for behaviour change, and the guidelines were not truly "owned" by the practitioners at whom they were aimed. In the second case the guidelines were part of a quality improvement process developed locally with encouragement, feedback and review. Thus, although the guidelines alone may not have led to changes in practice and outcome they were an essential ingredient.

Pilot testing, evaluation and revision

In general, better mechanisms are needed for designing, testing, implementing and revising practice policies (see Table 1). The US Agency for Health Care Policy and Research has suggested that standards and guidelines may undergo pilot testing before or while they are being disseminated.[3] We suggest that testing proposed guidelines under field conditions should be the rule rather than the exception. Revisions to improve their clarity and comprehensiveness would then be followed by repeat testing before dissemination of the final product. Poorly designed guidelines may create inconvenience, costs and risks to practitioners and patients alike. To promulgate guidelines in the absence of such testing may be akin to promoting the use of a promising new drug without adequate clinical testing.

The assessment of proposed guidelines or standards before their wide dissemination should address several factors.

- Acceptability: The guidelines should be acceptable to the intended audience (primary care physicians, specialists, nurses, allied health care professionals and so on).
- Comprehensibility: The guidelines should be assessed for their clarity and the extent to which practitioners understand them.
- Applicability: The guidelines must cover frequently encountered complicating factors and must be sufficiently flexible to allow application to the typical range of patients encountered. There may be tension between the ease of use and the comprehensiveness of the guidelines.
- Practicality: This includes anticipating changes in patient status, planning for results after the first stage of implementation (e.g., specifying basic treatment pathways after cholesterol screening) and considering the limitations in facilities available to practitioners.

Once the guidelines have been disseminated, with or without corresponding incentives and quality assurance mechanisms, further evaluation is warranted. The sponsors should seek empiric verification that the recommendations are reliably adopted and have led to changes in the process of care (because of compliance by practitioner and client) and in the expected health outcomes of patients.

Last, unless mechanisms exist to update guidelines in response to changes in knowledge, well-accepted guidelines may actually impede responses to new information; that is, one risk of "cookbook"

Table 1: The policy iteration cycle

Step	Component
Problem definition	Problem identified (e.g., excessive small-area variation, poor quality of care for common problem) Needs assessment done (relevant to providers, consumers and payers)
Setting of goals	Target group and parameters for success identified
Development	Group judgement process chosen; explicit criteria set for evidence and inference when possible
Pilot testing	Acceptability, comprehensibility, applicability and practicality tested in the field
Refinement	Identified barriers dealt with; implementation plan developed
Dissemination and implementation	Policies implemented according to plan
Assessment	Compliance and concordance with goals (based on process or outcome) assessed
Revision	"Plan–do–check–act" cycle iterated in response to findings of assessment; policies modified in response to new needs or information

medicine may be a loss in practitioners' ability and willingness to concoct their own recipes when new ingredients become available.

Conclusions

We can expect that increasing numbers of policies will be produced in response to the continued pressures on the health care system. By 1989 more than 26 US physician organizations had produced more than 700 practice standards and guidelines, and additional organizations had initiated plans to develop them.[37] This explosive growth in practice policies now parallels the growth in information that they were designed to address.

Practice policies deserve the same degree of rigorous consideration that is brought to bear on the development of the evidence they are based on. A clear statement of policy objectives and an explicit approach to their development, dissemination, implementation and evaluation is required (Table 1). This has yet to be fully realized, despite the innumerable policies tendered to date.

Practice policies are not a panacea for variations in practice and outcomes, uneven quality of care and the medical information explosion. Policies are attended by their own problems. National policies must be developed with a view to local needs and exigencies, and local programs, in turn, are crucial if policies are to be implemented successfully.

References

1. Haynes RB: Loose connections between peer-reviewed clinical journals and clinical practice. *Ann Intern Med* 1990; 113: 724-728
2. Wennberg JE, Freeman JL, Culp WJ: Are hospital services rationed in New Haven or over-utilised in Boston? *Lancet* 1987; 1: 1185-1189
3. *Clinical Guideline Development*, Agency for Health Care and Policy Research (program note), US Dept of Health and Human Services, Washington, 1990
4. *Webster's Ninth New Collegiate Dictionary*, Thomas Allen and Son, Markham, Ont, 1989: 910
5. Donabedian A: Criteria, norms and standards of quality: What do they mean? *Am J Public Health* 1981; 71: 409-412
6. The Expert Panel: Report of the National Cholesterol Education Program Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. *Arch Intern Med* 1988; 148: 36-69
7. Brook RH: Practice guidelines and practicing medicine: Are they compatible? *JAMA* 1989; 262: 3027-3030
8. Brook RH, Chassin MR, Fink A et al: A method for detailed assessment of the appropriateness of medical technologies. *Int J Tech Assess Health Care* 1986; 2: 53-64
9. Park RE, Fink A, Brook RH et al: Physician ratings of appropriate indications for six medical and surgical procedures. *Am J Public Health* 1986; 76: 766-772
10. Chassin M, Kosecoff J, Park RE et al: Does inappropriate use explain geographic variations in the use of health care services? *JAMA* 1987; 258: 2533-2537
11. Naylor CD, Baigrie RS, Goldman BS, Basinski A, Revascularization Panel and Consensus Methods Group: Assessment of priority for coronary revascularisation procedures. *Lancet* 1990; 335: 1070-1073
12. McIsaac W, Naylor CD, Basinski A: Mismatch of treatment intensity and coronary risk under the National Cholesterol Education Program guidelines. *J Gen Intern Med* 1991; 6: 518-523
13. Browner WS: Estimating the impact of risk factor modification programs. *Am J Epidemiol* 1986; 123: 143-153
14. Linton AL, Peachey DK: Guidelines for medical practice: 1. The reasons why. *Can Med Assoc J* 1990; 143: 485-490
15. Peachey DK, Linton AL: Guidelines for medical practice: 2. A possible strategy. *Can Med Assoc J* 1990; 143: 629-632
16. Kottke TE, Brekke ML: Cholesterol policy: What should we do? How should we decide [E]? *J Clin Epidemiol* 1990; 43: 1023-1027
17. Eddy DM: Clinical decision making: from theory to practice. Practice policies — What are they? *JAMA* 1990; 263: 877-878, 880
18. Idem: Practice policies: Where do they come from? Ibid: 1265, 1269, 1272
19. Idem: Clinical decision making: from theory to practice. Practice policies — guidelines for methods. Ibid: 1839-1841
20. Idem: Clinical decision making: from theory to practice. Guidelines for policy statements: the explicit approach. Ibid: 2239-2240, 2243
21. Idem: Clinical decision making: from theory to practice. Designing a practice policy: standards, guidelines, and options. Ibid: 3077, 3081, 3084
22. Canadian Task Force on the Periodic Health Examination: The periodic health examination. *Can Med Assoc J* 1979; 121: 1193-1254
23. US Preventive Services Task Force: *Guide to Clinical Preventive Services: an Assessment of the Effectiveness of 169 Interventions*, Williams & Wilkins, Baltimore, 1989
24. *Oxford English Dictionary*, compact ed, Oxford U Pr, Oxford, 1971
25. McGlynn EA, Kosecoff J, Brook RH: Format and conduct of consensus development conferences: a multination comparison. In Goodman C, Baratz R (eds): *Improving Consensus Development for Health Technology Assessment: an International Perspective*, Natl Acad Pr, Washington, 1990: 34-68
26. Relman AS: Assessment and accountability: the third revolution in medical care [E]. *N Engl J Med* 1988; 319: 1220-1222
27. Berwick DM, Godfrey AB, Roessner J: *Curing Health Care*, Jossey-Bass, San Francisco, 1990: 4, 41
28. Audet AM, Greenfield S, Field M: Medical practice guidelines: current activities and future directions. *Ann Intern Med* 1990; 113: 709-714
29. Tannenbaum TN, Sampalis JS, Battista RN et al: Early detection and treatment of hyperlipidemia: physician practices in Canada. *Can Med Assoc J* 1990; 143: 875-881
30. Hill MN, Weisman CS: Physicians' perceptions of consensus reports. *Int J Tech Assess Health Care* 1991; 7: 30-41
31. McGuire LB: A long run for a short jump: understanding clinical guidelines. *Ann Intern Med* 1989; 113: 705-708
32. Athanasoulis CA, Thrall JH: Standards of radiology practice: an approach to development. *Radiology* 1989; 173: 613-614
33. *Independent Health Facilities Act, 1989*, SO 1989, c 59
34. Wachtel TJ, O'Sullivan P: Practice guidelines to reduce testing in the hospital. *J Gen Intern Med* 1990; 5: 335-341
35. Lomas J, Anderson GM, Dominick-Pierre K et al: Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. *N Engl J Med* 1989; 321: 1306-1311
36. Myers SA, Gleicher N: A successful program to lower cesarean-section rates. *N Engl J Med* 1988; 319: 1511-1516
37. Kelly JT, Swartwout JE: Commentary: development of practice parameters by physician organizations. *QRB* 1990; 16: 54-57