

1

Introduction

Everywhere the old order changeth, and happy those who can change with it.

Sir William Osler, 1895

Guidelines for the provision of clinical care have been linked in recent years to almost every major problem and proposed solution on the American health policy agenda. Practice guidelines have been tied in some way, by some individual or organization, to costs, quality, access, patient empowerment, professional autonomy, medical liability, rationing, competition, benefit design, utilization variation, bureaucratic micromanagement of health care, and more. The concept has acted as a magnet for the hopes and frustrations of practitioners, patients, payers, researchers, and policy makers.

The broadest hopes of all parties are that practice guidelines will raise the quality of care and improve both the real and the perceived value obtained for health care spending. Beyond such widely held aspirations, individual groups differ in the emphasis they place on other narrower objectives. For example, administrators, regulators, and purchasers tend to stress cost control and reduced variation in practice patterns much more than physicians do. Practitioner groups tend to emphasize professionally developed guidelines as a means to maintain autonomy and to free professional decision making from external micromanagement. Consumer and patient advocates focus on guidelines to inform patients' decisions, clarify patient preferences, and strengthen patient autonomy.

Each group that has positive objectives for practice guidelines also fears their misuse. Their fears are essentially the obverse of their hopes--less sensitivity to patient needs, poorer outcomes, increased costs, lower quality, reduced autonomy or "cookbook" medicine, more bureaucracy, and

greater inequity in resource use. In particular, many physicians, especially those longer in practice, see guidelines as a challenge to clinical judgment and resist them as a threat to the most fundamental element of professional autonomy.

Recent public attention notwithstanding, guidelines are not new. Professional organizations have been developing guidelines for at least half a century, and recommendations about appropriate care can be found in ancient writings (Chassin, 1988). What is new is the emphasis on systematic, evidence-based guidelines and the interest in processes, structures, and incentives that support the effective use and evaluation of such guidelines.

Carefully developed guidelines for clinical practice can become part of the fabric of health care in this country and serve as important tools for many desirable changes. Their potential reach extends from improving the quality of clinical care (and its measurement) to helping to reduce the financial costs of inappropriate, unnecessary, or dangerous care. Practice guidelines are among the building blocks for informed patient decision making and rational social judgments about what care should be covered by public and private health benefit plans.

To the extent that guidelines provide well-argued translations of scientific research and expert judgment framed as statements about appropriate care, they will be more readily accepted by many kinds of decision makers. Such acceptance in the domains of physician practice, health education, quality assurance, medical liability, cost management, and elsewhere will provide mutually reinforcing support for the application and improvement of practice guidelines. Guidelines are not the solution to the country's health care problems, but they do have a significant, useful role to play.

As tools and building blocks for positive change, guidelines need to be understood and encouraged in context. That context includes powerful economic interests; changing and sometimes conflicting attitudes about professional and patient autonomy; policy making and implementing institutions that are intensely stressed and sometimes incapacitated; and scientific research that simultaneously expands both knowledge and uncertainty. Above all, the context in which guidelines will be used includes the complex, intimate relationship between individual patients and practitioners who are trying to protect health, manage illness, and preserve dignity under conditions that range from routine to desperate.

Also relevant are other strategies or forces for change that have their own challenges and uneven pace. Better clinical and outcomes research cannot produce results quickly, but the knowledge such studies generate will both strengthen guidelines over the longer term and build structures and processes for more constructive monitoring and feedback of information on performance to clinicians, managers, and others. Generational change, which obviously takes time, should lead to some greater acceptance of standardized, science-based guidelines as it brings to the fore practitioners,

administrators, and patients who have been socialized in an era of growing resource constraints, oversight, and technological and organizational complexity. If new quality improvement models can be successfully applied and sustained, they may provide a more positive environment for evidence-based practice guidelines. Technological advances in information systems may help guidelines on all fronts—in development, application, evaluation, and revision. The pace and nature of developments in each of these areas will influence the content, acceptability, and impact of practice guidelines.

PURPOSE OF THE REPORT

The examination undertaken and reported here has had two objectives: first, to encourage constructive expectations for guidelines and, second, to promote the kind of care and rigor in their development, application, evaluation, and revision that will help such expectations be realized. The charge to the Institute of Medicine (IOM) study committee had three parts: (1) describe existing initiatives to develop, use, evaluate, and improve guidelines for clinical practice, (2) assess the strengths and limitations of these initiatives, and (3) based on these assessments, recommend a conceptual and practical framework for the future development, use, and evaluation of guidelines. This framework could include whatever new public and private institutional arrangements seemed to be needed and feasible.

The committee has built on the work of other groups including previous IOM committees. In particular, its starting point was the 1990 IOM report, *Clinical Practice Guidelines: Directions for a New Program* (IOM, 1990c). That report provided advice to the Agency for Health Care Policy and Research (AHCPR) and its Forum for Quality and Effectiveness in Health Care. Its recommendations focused on guideline development, stressing the need for (1) systematic, science-based processes for developing guidelines, (2) careful documentation of the assumptions, evidence, and rationale for recommendations, and (3) explicit projections of the health and cost outcomes expected from the use of particular services or procedures. (This report often draws on the earlier report without specific reference.)

To conduct this more comprehensive examination of practice guidelines, the IOM appointed a committee of experts in the spring of 1990. Appointments included experts in medical and nursing practice, clinical and health services research, research methodology and program evaluation, medical informatics, and health care policy, financing, and administration. Approximately half the committee participated in the first IOM project for AHCPR.

IOM staff organized and conducted five meetings of the committee between June 1990 and September 1991. Other study activities and sources of information included several staff, committee, and commissioned papers, a public hearing, site visits, and focus groups. In addition, the committee established a liaison panel representing major organizations involved in the

development, use, and evaluation of practice guidelines. It also created a specialty society panel to assist communication with these groups. (See [Appendix C](#) for rosters of these panels.) This report was drafted, circulated to the committee for comment, revised, and then submitted for review in accordance with IOM and National Research Council (NRC) report review policies. The report was revised again based on the NRC review, and this document constitutes the committee's final report.

The primary audiences for this study are public and private policy makers, specialty society leaders, and managers of institutions or organizations that may support the application of practice guidelines. Others who are not likely to read an IOM report firsthand may nevertheless learn and benefit from the study as it is discussed in journals, conferences, and similar venues.

Throughout this report, implementation of practice guidelines is a particular focus. Policy makers, researchers, guidelines developers, and others have thus far paid more systematic attention to guidelines development than to guidelines implementation or evaluation. In contrast to that emphasis, in this document even the chapters on development of guidelines emphasize how the content of guidelines and the way they are developed may affect their use.

WHAT ARE PRACTICE GUIDELINES?

Definitions of Key Terms

Definitions for the word *guidelines* abound, as do other terms that different organizations or individuals prefer to use instead of guidelines (IOM, 1990c). Sometimes the term *practice guideline* serves as an umbrella label for practice standards, protocols, parameters, algorithms, and various other types of statements about appropriate clinical care; at other times, sharp distinctions are drawn. Debate about terminology reflects, in part, controversy and disagreement about the uses of guidelines and related materials.

This report, like its predecessor, uses the term *practice guideline*, largely because it is the term most generally used. For example, the 1989 legislation that created AHCPR and the Forum employed the term.¹ Prac

¹ Two terms found in OBRA 89 (the Omnibus Budget Reconciliation Act of 1989)-standards of quality and performance measures-are not used here. The first term has quite different and even contradictory uses. The 1990 IOM report to AHCPR defined standards of quality as "authoritative statements of (1) minimum levels of acceptable performance or results, (2) excellent levels of performance or results, or (3) the range of acceptable performance or results." Statements described as standards should clearly indicate whether they articulate minimums, maximums, or ranges of quality. The term *performance measures* has no clear professional usage, and the 1990 report defined them provisionally as "methods or instruments to estimate or monitor the extent to which the

tice guideline also will be the main entry term for the reference subject headings used by the National Library of Medicine to index the literature on health services research (Peri Schuyler, National Library of Medicine, personal communication, May 20, 1991).

As defined in the IOM's 1990 report, practice guidelines are "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances"² Medical review criteria are "systematically developed statements that can be used to assess the appropriateness of specific health care decisions, services, and outcomes."

Practice guidelines focus, in the first instance, on assisting patients and practitioners in making decisions, but this defining characteristic does not and should not preclude their use for other purposes including quality assurance and payment policy making. Conversely, medical review criteria and related tools emphasize the evaluation of health care decisions, actions, and outcomes, but they should and do build on guidelines and may in some cases be virtually identical.³

Practice guidelines are *not* synonymous with the reimbursement or coverage policies of Medicare and other health insurance plans, which traditionally have excluded some items from coverage (for example, immunizations and blood products) for reasons unrelated to the appropriateness of the service. Reimbursement and coverage policies certainly may be informed by practice guidelines, but this report attempts to distinguish between the two.

Although the IOM definition of clinical practice guidelines emphasizes those aimed at specific clinical problems such as diabetes, some apply to very broad ranges of clinical problems, patients, and services. For example, so-called universal precautions seek to control human immunodeficiency virus (HIV) and other infections by requiring that certain practices (such as using gloves and discarding needles in special containers) be followed for

² This report often refers, first, to "sets of guidelines," which present a series of statements about appropriate care and, second, to "guidelines documents," which may include short statements of recommendations, longer presentations summarizing methods and evidence, and very long documents describing methods, evidence, rationales, and other issues in great depth.

³ One committee member strongly objected to the distinction between guidelines and review criteria, arguing that "there should not be one iota of difference between a good guideline intended for [practitioners] and a medical review criterion intended to assess care; they are different uses of the same clinical statement." The committee felt, on the whole, that distinguishing guidelines aimed at clinicians or patients from review criteria aimed at assessing care was useful even though the latter may and should draw on the former. In fact, given the importance accorded to quality assessment and cost containment objectives, some organizations may choose the development of review criteria as their starting point; however, the result may be statements that are presented in

all patients whether or not they are known to be infected. Guidelines for informed consent policies likewise apply quite generally. Broad guidelines, which are frequently adopted as institutional policies, sometimes in response to accreditation standards set by the Joint Commission on Accreditation of Health care Organizations or other bodies, may reflect difficult ethical, legal, and management issues as well as clinical concerns. Recent American Medical Association (AMA) guidelines on the use of "Do Not Resuscitate" orders are a case in point; they were issued by the AMA's Council on Ethical and Judicial Affairs (1991b). Later in this report, the committee recommends the development of general guidelines on information for patient decision making.

A final note on terminology: the definitions of guidelines and review criteria refer to *appropriate care*, a term that also presents definitional problems. Sometimes it is used as a synonym for required care; at other times it seems to be viewed (consistent with dictionary usage) as care that is suitable or proper but not always necessary, absolutely required, or essential. For purposes of this report, appropriate care is conceptually defined as care for which "the expected health benefit [exceeds] the expected negative consequences by a sufficient margin" that the care is worth providing (Park et al., 1986, p. 6). (Concepts of necessary, appropriate, or minimum care resurface in [Chapter 6](#).)

Terminological disagreements undoubtedly will continue, and rigid distinctions could sabotage some productive discussions. The field of clinical practice guidelines is still developing, and different terminology may prove more functional in the future.

Operationalizing this report's conceptual definition of guidelines is an exercise fraught with difficulties, both technical and normative. The general strategy urged in this report calls for developers of guidelines to state how compelling is the case for a particular course of care based on the strength of the evidence, the strength of professional judgment, and the importance of the benefits. If the case is clearly stated, others will have information and a model for evidence-based decision making that they may use to reach different judgments.

Desirable Attributes of Guidelines

The above definitions attempt to identify essential characteristics and are not intended to describe the qualities that good practice guidelines and review criteria should have in every use to which they can be put. Thus, the first IOM committee on guidelines identified and discussed eight desirable attributes of guidelines for clinical practice. [Chapter 5](#) of this report presents desirable attributes for medical review criteria. These two sets of attributes must be viewed as statements of aspirations, which are intended to encourage developers of guidelines and review criteria to improve their

processes and products; they are not meant as vehicles for destructive criticism.

This committee has made one modification in the list of eight attributes proposed in 1990.⁴ Under the attribute of validity, it adds that every set of guidelines should be accompanied by (1) a statement of the strength of the evidence and the expert judgment behind the guidelines and (2) projections of the relevant health and cost outcomes of alternative courses of care. Assessments of relevant health outcomes should consider patient perceptions and preferences. The attributes as amended are presented in [Table 1-1](#).

The committee had two primary reasons for these amendments. First, those citing or using the first IOM report have tended to stress the formal list of attributes without mentioning the elements of validity that were identified in the accompanying text. Second, evidence, outcomes, and patient decision making are emphasized in this second report, particularly in the discussion of ethics, costs, and informed consent; the amendments to the list of the attributes reflect this emphasis.

Four of the eight attributes relate to the content of guidelines: validity, reliability, clinical applicability, and clinical flexibility. Four others relate to the process of guideline development or the presentation of guidelines: clarity, multidisciplinary process, scheduled review, and documentation. Each affects the likelihood that guidelines will be perceived as credible and usable, and the probability that they will, if used, help achieve desired health outcomes. Collectively, these attributes tend to be what distinguishes systematically developed practice guidelines from general textbook knowledge, although the boundaries between these (and other) kinds of information or recommendations are not well defined. Because the IOM and AHCPH recognized that it would be useful but difficult to employ these attributes to assess practice guidelines, one objective of this study has been to develop a practical instrument to guide such assessments. The results are described in [Chapter 8](#) and [Appendix B](#).

Elements of Analysis

The above attributes imply a challenging analytic strategy for developers of practice guidelines that reflects a rigorous scientific process—"a rigorous and orderly asking and answering of questions," in the words of one reviewer of this report.⁵ The components of such a strategy can be briefly summarized as

⁴ The attributes are discussed at some length in [Chapter 3](#) of the first IOM report, and readers are urged to consult that text for a fuller understanding of each attribute.

⁵ More detailed discussions of analytic strategies and steps for developers of

TABLE 1-1 Desirable Attributes of Clinical Practice Guidelines

Attribute	Explanation
VALIDITY	Practice guidelines are valid if, when followed, they lead to the health and cost outcomes projected for them. A prospective assessment of validity will consider the substance and quality of the evidence cited, the means used to evaluate the evidence, and the relationship between the evidence and recommendations.
Strength of Evidence	Practice guidelines should be accompanied by descriptions of the strength of the evidence and the expert judgment behind them.
Estimated Outcomes	Practice guidelines should be accompanied by estimates of the health and cost outcomes expected from the interventions in question, compared with alternative practices. Assessments of relevant health outcomes will consider patient perceptions and preferences.
RELIABILITY/REPRODUCIBILITY	Practice guidelines are reproducible and reliable (1) if—given the same evidence and methods for guidelines development—another set of experts produces essentially the same statements and (2) if—given the same clinical circumstances—the guidelines are interpreted and applied consistently by practitioners (or other appropriate parties).
CLINICAL APPLICABILITY	Practice guidelines should be as inclusive of appropriately defined patient populations as evidence and expert judgment permit, and they should explicitly state the population(s) to which statements apply.
CLINICAL FLEXIBILITY	Practice guidelines should identify the specifically known or generally expected exceptions to their recommendations and discuss how patient preferences are to be identified and considered.
CLARITY	Practice guidelines must use unambiguous language, define terms precisely, and use logical and easy-to-follow modes of presentation.
MULTIDISCIPLINARY PROCESS	Practice guidelines must be developed by a process that includes participation by representatives of key affected groups. Participation may include serving on panels that develop guidelines, providing evidence and viewpoints to the panels, and reviewing draft guidelines.
SCHEDULED REVIEW	Practice guidelines must include statements about when they should be reviewed to determine whether revisions are warranted, given new clinical evidence or professional consensus (or the lack of it).
DOCUMENTATION	The procedures followed in developing guidelines, the participants involved, the evidence used, the assumptions and rationales accepted, and the analytic methods employed must be meticulously documented and described.

- formulation of the problem (for example, the clinical condition to be considered, the key issues to be addressed, and the relevant alternative courses of care to be examined, which may include "watchful waiting");
- identification and assessment of the evidence from clinical trials, case-control studies, and other sources to determine where evidence is weak, missing, or in dispute;
- projection and comparison of health benefits and harms (including how they are perceived by patients) associated with alternative courses of care;
- projection of net costs associated with achieving the benefits of alternative courses of care;⁶
- judgment of the strength of the evidence (considering key areas of scientific uncertainty and theoretical dispute), the relative importance of the projected benefits and risks (again with patient perspectives considered), and—overall—how compelling is the case for particular interventions;
- formulation of clear statements about alternative courses of care, accompanied by full disclosure of the participants, methods, evidence, and criteria used to arrive at these statements; and
- review and critique of all these elements by methodologists, clinicians, and other relevant parties not involved in the original process.

This framework is only a brief summary of the strategy that is elaborated on at various points throughout the report. The analytic steps identified above can be managed by a single organization. Alternatively, different parties may contribute to the process. Today, for example, most professional societies do not consider costs or patient preferences. Others can add these steps later, although such additions will be more difficult if the initial work has not anticipated the questions to be asked in these later analyses.

In addition to following this analytic strategy in developing guidelines, developers should seek to present their work to clinicians in ways that reflect the rigor of this approach and its emphasis on reasoning and critical analysis. The product of the process should not be perceived solely as information but more broadly as an explication of the thinking processes that should be used in evaluating and applying that information. If guidelines are perceived only as information, they may very well be used (or rejected) as the "cookbooks" that many physicians decry.

⁶ In addition, comparisons may usefully involve different clinical problems as well as different approaches to the same problem. For example, the cost-effectiveness of screening for hypertension (\$16,280 per quality-adjusted life year—or QALY—for asymptomatic men aged 60) has been compared not only with other heart disease screening but also with treatment of heart disease (such as surgery for left main coronary

Guidelines presented to patients will necessarily be simpler than those presented to physicians, but they, too, should try to emphasize responsible decision making and not just cut-and-dried advice or information. As guidelines are developed that are more sensitive to variations in patient preferences and the role of patients in making decisions, the initial formulation of guidelines is likely to make their translation into patient-usable forms easier. These initial formulations should clearly describe the possible outcomes of alternative management strategies in terms that are relevant to patients, discuss what is known about variations in patient preferences for different outcomes, and note points at which patient choices among alternatives should be requested.

The analytic framework presented above represents an ideal. Making progress toward this ideal will take time. Some, perhaps most or all, guidelines will not fully reach it. The committee recognizes this to be the case but, at the same time, emphasizes the importance of keeping the ideal in mind and making a serious and persistent effort to achieve it. In addition, the committee urges research methodologists and others to work to improve (and, when possible, simplify) the procedures and tools for analyzing evidence, reaching responsible group judgments, and otherwise arriving at sound recommendations for care (see [Chapter 7](#) for further discussion).

State of the Evidence

In developing guidelines, conclusions backed by scientific evidence should take precedence over statements based on subjective judgments. When the empirical evidence has important limitations (as will typically be the case) or when experts reach conclusions that are not consistent with the evidence, the limits of the evidence should be clearly described and the rationale for departing from it should be explained. When expert judgment proceeds in the absence of direct empirical evidence about a particular clinical practice, the general scientific reasoning or normative (ethical, professional) principles supporting the expert judgments should be described. Statements about the importance of particular benefits and harms will reflect both empirical analyses and value judgments; [Chapter 6](#) returns to this point.

For users of guidelines, this kind of argumentation, reasoning, and documentation can help in sorting out conflicting claims, considering how guidelines should or should not be adjusted to local circumstances, and independently evaluating the claims made for guidelines. That the relationship of evidence to recommendations cannot be taken for granted is illustrated by an analysis of recommendations on dietary cholesterol that found virtually all the cited references to be irrelevant or in conflict with the recommendations (Reiser, 1984; see also Eddy and Billings, 1988).

To the extent that guidelines move toward statements and arguments

such as those outlined above, they will identify how compelling is the case for particular services or courses of care under particular clinical circumstances. They will distinguish care that is strongly or moderately supported (or contraindicated) by strong scientific evidence and consensus, care that is supported chiefly by consensus without any direct research backing, and care about which experts differ in the face of mixed or absent evidence. Along these lines, Eddy (1990e) has divided statements about appropriate care—what he calls practice policies—into three categories depending the clarity of the evidence about outcomes and the importance of the outcomes to patients. When the case for a particular course of care is very strong, *standards* can be delineated for care that is to be provided or recommended to patients with only rare circumstances justifying exceptions. When the case is somewhat less compelling, *guidelines* (used by Eddy in a narrower sense than in this report) can be defined for courses of care to be provided or recommended in most cases but with more exceptions allowed than are warranted for standards. *Options* note that different courses exist and that evidence does not warrant specific recommendations.⁷

By the term *strong evidence*, the committee refers to (1) the characteristics of the evidence itself (for example, whether it shows a strong effect, no effect, an inconclusive effect, or something in between) and (2) the qualities of the process for generating that evidence. Formal hypothesis-testing processes range in strength from experimental to quasi-experimental to nonexperimental. However, a strong research design that is improperly executed may provide poorer evidence than a weaker but properly executed design. Single case reports and case series do not test hypotheses but do provide relatively weak forms of empirical evidence. Formal methods of generating expert consensus yield evidence of what clinicians believe about a particular form of care, based on their experience and their assessment of such evidence as does exist; statements of consensus may provide useful guidance but do not constitute clinical evidence as the term is used here. Problems involved in rating and combining evidence are revisited in [Chapter 7](#).

Inevitably, given the state of scientific knowledge, many courses of care will not be supported by good evidence. [Table 1-2](#) presents a purely hypothetical (but the committee believes plausible) illustration of how evidence and consensus might be distributed across the entire range of health care services. It is based on an example offered by one committee member

⁷ In a similar vein, the American College of Cardiology has distinguished three broad classes of guidelines: (1) general agreement exists that the service/technology is

Copyright 2004 © National Academy of Sciences. All rights reserved. service/technology is not appropriate;

Unless otherwise indicated in individual titles, this PDF file is provided by the National Academies Press (www.nap.edu) for research

purposes are copyrighted by the National Academy of Sciences. Distribution, posting, or copying is strictly prohibited without

written permission of the NAP.

Generated for eesandi@infovia.com.ar on Mon Feb 28 08:33:14 2005

About this PDF file: This new digital representation of the original work has been recomposed from XML files created from the original paper book, not from the original typesetting files. Page breaks are true to the original; line lengths, word breaks, heading styles, and other typesetting-specific formatting, however, cannot be retained, and some typographic errors may have been accidentally inserted. Please use the print version of this publication as the authoritative version for attribution.

as a means of clarifying the committee's understanding of the range of good science and evidence in today's world.

TABLE 1-2 Hypothetical Distribution of Evidence and Consensus for All Health Services and Patient Management Strategies

Strength of Evidence	Strength of Consensus	Percentage of All Services
++	++	2
++	+	2
++	-	0
+	++	20
+	+	25
+	-	0
-	++	20
-	+	25
-	-	6

NOTE: ++ strong; + modest; - very weak or none.

A guideline having strong evidence and strong consensus is the U.S. Preventive Services Task Force (USPSTF, 1989) recommendation that erythromycin ophthalmic ointment be used for all babies as soon as possible after birth to prevent gonococcal ophthalmia neonatorum infection.⁸ In the category of little evidence but general consensus is the widely accepted, easily remembered blood "transfusion trigger" (hemoglobin levels of less than 10 grams per deciliter). There is no rigorous (and little nonrigorous) clinical research that evaluates patient outcomes for transfusions at this or other levels (Welch et al., 1991). Clearly, however, for a physician faced with a woman bleeding to death from a ruptured ectopic pregnancy or some similar emergency, the absence of research on specific thresholds for transfusion cannot be a counsel for inaction.⁹

⁸ As rated using a scheme formulated by the USPSTF, the statement was based on evidence rated "I" (drawn from at least one properly designed randomized clinical trial) and strength of recommendations rated "A" (good evidence for the recommendation).

⁹ The situation portrayed in Table 1-2 would be even more stark if the universe of care were defined in very detailed clinical terms. This point can be illustrated with one currently applied guideline for hysterectomy (Mark Chassin, Value Health Sciences, personal communication, September 19, 1991). This guideline (part of a larger set of guidelines) states that hysterectomy is inappropriate for women with all these characteristics: under age 30, no children, an expressed desire for no future pregnancy, mild dysfunctional bleeding (defined objectively and clinically), one dilatation and curettage in the previous 12 months, and no trial of hormone therapy. At this level of clinical detail, there are no outcome data, no functional status data indicating whether women experience the condition as impinging on their daily lives, no data on

Those who develop guidelines can highlight areas for which evidence exists, for which it is missing, and for which it is flawed. This process will identify specific holes in research—for example, the absence of work on appropriate intervals for blood pressure screening. It will also provide researchers and research funders with a helpful picture of the gaps in whole categories of research questions (for example, testing intervals in general).

Any discussion of the state of scientific evidence must also note the challenge posed by the rapid advance of clinical research (McGuire, 1990). Months of effort may be rendered largely or partly irrelevant by new information; for example, follow-up results may challenge earlier findings, or convincing findings from clinical trials may arise unexpectedly. This fact of life underscores the importance of processes for updating guidelines and for disseminating important contradictory research findings. However, constantly changing guidelines (based on changing data and consensus) are cited as one reason for an unsuccessful Swedish initiative to develop and apply practice guidelines in the 1970s (Little, 1990).

Inconsistent or Conflicting Guidelines

Guidelines are one means of clarifying acceptable and unacceptable variation in medical practice. Nevertheless, that clarification itself has limits that may lead different groups to different and even inconsistent guidelines. Weak evidence is still weak evidence, although the processes discussed in this report should allow the best use of whatever evidence is available. Moreover, differences of opinion (and, thus, differences in guidelines) can be expected about such matters as whether a research design flaw "matters" or whether differences in the results of two treatment alternatives are "clinically important" or only "statistically significant." In addition, individuals and groups will vary in their values and tolerance of risk. (For an interesting illustration of these factors at work in the debate over childhood cholesterol screening, see Newman et al., 1990 and Resnicow et al., 1991).

When evidence is limited or nonexistent, developers of guidelines have used different strategies for making recommendations (Hayward et al., 1991). Some offer recommendations; others do not. In any case, this committee calls for guidelines to explain the rationale for the presence or absence of a recommendation and to describe how compelling is the case for alternative approaches to particular clinical problems. Guidelines that do this can build a more credible and more powerful base for decision making by patients, practitioners, and others. Ways of dealing with inconsistent or conflicting guidelines are discussed further in [Chapter 7](#).

Types of Guidelines

Guidelines and medical review criteria can be categorized along many dimensions. As illustrated in [Appendix A](#), guidelines and guidelines-like materials may vary in five main ways:

- **Clinical orientation.** Some guidelines deal with clinical conditions or problems (for example, throat infections in children), whereas others describe the indications for using procedures or services (for example, tonsillectomy).¹⁰
- **Clinical purpose.** Guidelines may address several broad kinds of health care interventions: (1) screening and primary prevention, (2) diagnosis, (3) treatment and management (including secondary prevention), and (4) rehabilitation.
- **Complexity.** Complexity is a function of many factors: the nature of the specific clinical conditions or technologies being dealt with; the extent and certainty of knowledge about the conditions or technologies; the options and interrelationships among options for managing the conditions; and the objectives, approaches, and skills of those developing the guidelines. As a case in point, the number of appropriateness criteria developed by researchers at the RAND Corporation ranges from 49 for cholecystectomy (Solomon et al., 1986) to 2,862 for colonoscopy (Kahn et al., 1986).
- **Format.** Format refers to how guidelines (particularly the statements about appropriate care rather than all the supporting documentation and rationales) are physically presented, whether in free text, through tables or other graphics, as algorithms, or by other means.
- **Intended users.** As noted elsewhere in this chapter and in this report, the sets of potential users are quite large and diverse; for purposes of the descriptions used in [Appendix A](#), the main categories are "practitioners" and "patients."

WHY ARE POLICY MAKERS INTERESTED IN GUIDELINES?

Some would explain the interest in practice guidelines shown by legislators, regulators, and purchasers of health care (as opposed to that of practitioners and patients) with a single phrase: out-of-control health care costs. If, despite nearly two decades of intensifying efforts to contain spending, health care costs had not been increasing substantially faster than costs in other sectors, most of the recent legislation, conferences, and other activities to promote guidelines probably would not have happened despite the

other concerns—most notably, quality improvement—that guidelines also address. Many interested parties may be disappointed if they think guidelines will not reduce costs (Bouxsein, 1988).

Although the importance of cost concerns as a stimulus for guidelines should not be understated, concerns about quality of care and risk management also figure prominently in the call for more and better practice guidelines. The attraction of guidelines also has had a political component. Guidelines were offered by and to physician groups as an acceptable, partial alternative to the specter of more stringent controls on Medicare payments for physician services (American Society of Internal Medicine [ASIM], 1989, 1990; Physician Payment Review Commission [PPRC], 1988, 1990; Kosterlitz, 1991). They were promoted as a selective approach that targets inappropriate or unnecessary care and relies on informed decision making by practitioners and patients rather than by far-removed officials.

More specifically, the growing interest in guidelines has been prompted by perceptions, first, that higher health care expenditures have brought only marginal health benefits and, second, that guidelines can help remedy this problem of "value." Virtually every major discussion of guidelines begins with a similar list of reasons for these perceptions (PPRC, 1988, 1989; IOM, 1989a; Billings, 1990; Leape, 1990; Hammons, 1991). The discussion generally proceeds as follows.

- Research demonstrates major **variations in physician practice patterns and utilization of health services** (Wennberg and Gittelsohn, 1973, 1982; Wennberg, 1984, 1991; Chassin et al., 1986a, 1987; R.E. Brown et al., 1989). The lowest level of use may not be the right level, but the variations raise troubling questions about the justification for these variations and their accompanying costs.
- Other research indicates considerable **inappropriate use of many services** including laboratory tests, diagnostic and surgical procedures, prescription medications, and inpatient hospital admissions and days of care (Brook et al., 1986; Eisenberg, 1986; Lohr et al., 1986; Chassin et al., 1987; Foxman et al., 1987). Estimated inappropriate use of care for selected services ranges from 10 percent to more than 30 percent; estimates of associated unnecessary expenditures vary widely. Many of the services studied, however, have been those particularly suspected of overuse; even for these services, some degree of underuse may also exist.
- In addition, much health care is characterized by considerable **uncertainty about the health outcomes** achieved by the use or nonuse of various services and procedures (Office of Technology Assessment [OTA], 1978; Dersimonian et al., 1982; Eddy, 1984; Wennberg, 1984; Eddy and Billings, 1988; Roper and Hackbarth, 1988; Roper et al., 1988; Brook, 1989, 1990; IOM, 1989b, 1990a,b,e,f,h). Clinical research documenting the effective

ness of many services does not exist, particularly at the level of very specific patient circumstances; thus, the value received for spending on these services is likewise unknown.

Whether the issue is unexplained variation, inappropriate care, or uncertain outcomes, many analysts come to similar conclusions. More research on outcomes and effectiveness of health care services is needed; more work should be done, using such research, to formulate specific guidelines for clinical practice; and more use of the resulting guidelines will help limit health care spending.

How are guidelines to limit health care costs? How are they to increase the perceived value of health care spending? The basic argument or hypothesis runs along these lines:

Scientific evidence and clinical judgment can be systematically combined to produce clinically valid, operational recommendations for appropriate care that can and will be used to persuade clinicians, patients, and others to change their practices in ways that lead to better health outcomes and lower health care costs.

Six formidable and often unrealistic assumptions or expectations lie behind this partly explicit and partly implicit causal model.

- First, scientific evidence of sufficient quantity and quality exists to serve as a foundation for guidelines.
- Second, programs to develop guidelines will be organized, funded, and effectively managed to produce a considerable volume of valid, usable statements about appropriate care for clinically and financially significant health conditions or technologies.
- Third, substantial numbers of clinicians, patients, and others will have the opportunity, the support, and the incentives to read, understand, accept, and use these statements in ways that change patterns of clinical practice, health behavior, or payment for health care services in desired directions.
- Fourth, such changes will be broad and intense enough to improve health outcomes.
- Fifth, on balance, the entire body of guidelines as actually developed and used will lead to more cost-controlling than cost-increasing behavior on the part of providers.
- Sixth, the body of guidelines will continually expand to cover new areas so that net rates of increase in health care costs and absolute levels of expenditures will be lower than they would otherwise be.

Unfortunately, these six expectations outstrip current capacities in several respects. For many clinical conditions and services, the science base is limited, and even when it is reasonably satisfactory, clinicians and analysts

may disagree in their interpretations of the evidence (recall [Table 1-2](#)). Developing guidelines based on systematic, evidence-based processes is expensive and time-consuming, and the volume of such efforts, though increasing, is still small in relation to the scope of clinical care. Moreover, despite the good intentions of many involved parties, much guideline development remains relatively unsystematic; the enterprise as a whole still lacks proven mechanisms for evaluating, improving, and targeting the development of guidelines. Psychological, economic, and other factors limit clinician and patient acceptance of and conformance with guidelines. Organizational systems for quality, cost, risk, and information management are not planned and structured to support awareness, acceptance, and use of credible practice guidelines. For the uninsured, underinsured, and others, indicated care may not be affordable or otherwise accessible.

Even if the first four of these expectations about the scope, quality, application, and health outcomes of guidelines were to be fulfilled, the committee regards as questionable the last two expectations about the cost consequences of change. As argued earlier, some guidelines undoubtedly will save money by reducing the use of inappropriate or unnecessary services; some will increase expenditures by encouraging more use of underutilized services; and some will shift costs from one type of service to another or from one payer to another. The net impact of guidelines on the rate of increase in total health care spending cannot be predicted with confidence, even if future priorities for guidelines development stress clinical conditions for which costly overuse of services is suspected.¹¹

Furthermore, the current system of delivering and financing care does not have incentives for economy and efficiency that are strong and consistent enough to capitalize fully on the opportunities for cost control that some guidelines present. New technology and other factors also will continue to exert upward pressure on total costs, as will policies to improve access to care for the uninsured and other disadvantaged groups.

In sum, guidelines for clinical practice are a promising but not a quick or sure strategy for improving and rationalizing the use of health services. The attention and resources now invested in guidelines could dissolve in the face of a collision between unrealistic hopes and limited immediate results. For guidelines to fulfill their potential, persistent commitment over the long term is required from both policy makers and health care professionals.

¹¹ Besides costs, other relevant factors in selecting topics for guideline development include the potential for an assessment to change health outcomes, the amount of use of a technology), and the

WHO USES GUIDELINES AND FOR WHAT?

The potential users of clinical practice guidelines constitute a diverse group. This report focuses on the primary users of guidelines: practitioners, patients and families, and health care institutions. Other users include those payers, health benefit plans, and public policy makers and regulators who may use guidelines in making specific decisions about what health care to reimburse, cover, or encourage and in evaluating the decisions, actions, or performance of the primary users of guidelines. In addition, some individuals and organizations act, in a sense, as "conduits," facilitating or promoting the use of guidelines without directly applying them to make decisions. Examples of such users include educators of many sorts and science writers and journalists who may facilitate discussion and dissemination.

Any single user of guidelines may employ them in various ways, and any particular set of guidelines may need to be presented in different ways for different users and uses. Five major purposes for guidelines, which are not mutually exclusive, are

1. assisting clinical decision making by patients and practitioners,
2. educating individuals or groups,
3. assessing and assuring the quality of care,
4. guiding allocation of resources for health care, and
5. reducing the risk of legal liability for negligent care.

The first and second uses may reflect a fairly straightforward application of guidelines; the third and fourth typically entail the translation of guidelines into medical review criteria and other evaluation tools. The fifth use may be a more indirect product of the other uses, although some guidelines have been developed with this use in mind.

The definition of guidelines used in this report highlights one crucial purpose: to assist individual practitioners and patients in making decisions about specific clinical problems. For example, a physician might use a guideline to assess medical management of a given condition versus expeditious surgical intervention before discussing risks, benefits, and options with the patient. Another physician might consult a guideline to determine the appropriate prescription medication to use, given that medical management of an illness is warranted. A patient may consult guidelines in deciding whether to seek specific screening services. A nurse might review a guideline in preparing a care plan for a homebound patient or nursing home resident; a nurse-practitioner might check a guideline, perhaps in the form of a protocol, to determine whether to treat a patient or refer the patient to a physician. These examples represent the central uses envisioned by most developers of guidelines.

Guidelines are also used for individual educational purposes. Physicians, nurses, and others may rely on guidelines to help patients and families understand clinical situations and available courses of action. Depending on the complexity of the guideline, its distribution, and patients' prior level of knowledge, patients might also use guidelines fairly directly in their own decision making. In addition to individual educational uses, many guidelines are employed in continuing medical education, public health campaigns, and other organized programs to educate broad categories of professionals, patients, or others about appropriate health care or behavior.

To assess and improve the quality of health care, organizations and individuals may refer to practice guidelines (and review criteria) for several reasons: to structure organizational procedures, to guide equipment purchases and hiring decisions, and to set and implement priorities for monitoring, feedback, and other efforts to assess and improve performance. For example, health care plans may check their records to determine how successful their practitioners have been in immunizing children or screening adults for particular problems; depending on the results, they may then try to improve their reminder systems, patient education efforts, or other aspects of plan operation.

Given the hope of many that guidelines can help control health care costs, it is not surprising that individuals, health care organizations, and public and private payers refer to practice guidelines in making decisions about resource use and in attempting to influence the decisions of others. For example, practitioners and institutions at financial risk from their participation in capitated, per-case, or other non-cost-based payment schemes may employ guidelines and review criteria to identify wasteful patterns of care, avoid expensive purchases of equipment with few approved indications for use, and forestall inappropriate referrals to specialist consultants. Public and private payers may use practice guidelines or review criteria to help them make broad decisions about whether to cover particular services (for example, pancreas transplants) or to precertify the appropriateness of specific services for particular patients (such as carotid endarterectomy for asymptomatic individuals). To the extent that guidelines can be used to help rationalize the provision of health services, demands for explicit rationing of useful care may be avoided or minimized.

Specialty societies, health care institutions, malpractice insurers, and even legislators have become acutely interested in how guidelines and related review criteria may reduce the exposure of practitioners and institutions to malpractice liability. Such a use of guidelines, although worth noting in its own right, can also serve both quality-and cost-management goals. For example, less inappropriate or dangerous care should improve the processes and outcomes of care and reduce the number of malpractice claims and judgments, thereby reducing litigation and compensation costs.

BASIC PROPOSITIONS

This report offers six broad propositions about the current state and future role of clinical practice guidelines. First, practice guidelines can be (and are being) formulated and used now to improve the quality and value of health services. Even if limited in scope, they are a positive step.

Second, although guidelines development is not a firmly established enterprise with well-tested methods and procedures, efforts to develop practice guidelines are widespread, growing, and diverse. The movement is likely to remain pluralistic and perhaps in some ways even competitive—a circumstance that offers opportunity for growth and progress as well as some risk of confusion and contradiction.

Third, a major challenge for practice guidelines is better follow-through. To capitalize on sound guidelines that constructively anticipate practical problems faced in real clinical situations, practitioners, policy makers, and others need better strategies and processes to ensure that guidelines are effectively implemented.

Fourth, science can contribute more effectively to useful knowledge-based guidelines by devising research strategies to evaluate the effectiveness of emerging and existing services that better reflect the conditions of actual practice. The gap needs to be reduced between, on the one hand, clinical research conducted on homogeneous populations within carefully controlled settings and, on the other hand, effective knowledge for those providing health care to heterogeneous populations in diverse settings.

Fifth, although effectiveness research and clinical practice guidelines can inform action and contribute to basic ethical debates over what constitutes an appropriate distribution of resources or an appropriate structure for health care delivery, they can not resolve those debates. Decisions depend on many other factors including political judgments, cultural norms, economic calculations, and the power of affected interests.

Sixth, expectations that practice guidelines will help control total health care spending should be restrained. Wider application of guidelines aimed at currently overused services will likely reduce some spending. In other cases, spending will shift from inappropriate to more appropriate care. At the same time, guidelines that focus on currently underused services may stimulate increased expenditures, particularly if strategies to improve access for those who are not now adequately insured are successful. Collectively, the result undoubtedly will be better value but not guaranteed net savings— particularly given the health care system's lack of strong, consistent incentives for efficient and economical behavior by practitioners, patients, and others.

These caveats notwithstanding, clinical practice guidelines have real potential to help clarify the knowledge base for clinical practice and to

improve the quality and effectiveness of medical care. Realization of that potential will depend on astute policy making and steady management, as well as on sound scientific evidence and clinical judgments.

REPORT ORGANIZATION

The next chapter of this report discusses current efforts to develop clinical practice guidelines. The emphasis is on national activities; local development and adaptation of guidelines are considered later. As a prologue to later chapters, it begins by stressing that what happens during the development process will influence the probability of successfully implementing guidelines for clinical practice.

Chapter 3 provides an overview of guidelines implementation, including the various factors that shape decisions about implementation strategies and affect their application. In an attempt to bring to this report some sense of the real world in which guidelines are—and are not—used, the chapter uses several hypothetical case studies to illustrate various implementation issues.

Chapter 4 examines the societal context and the philosophical and strategic considerations that may influence efforts to bring guidelines into use. It also considers specifically the roles of education and of information and decision support systems. **Chapter 5** then discusses how systems to manage quality, costs, medical liability, and information may support and be supported by clinical practice guidelines. It also proposes desirable attributes of medical review criteria.

Chapter 6 considers the pervasive issue of health care costs and some of the related ethical, political, and technical controversies about how guidelines should be developed and used. It includes an examination of issues in cost and cost-effectiveness analysis and consideration of informed consent and minimum standards of care.

Using the discussion in the preceding three chapters as a base, **Chapter 7** returns to the procedures and methods for developing practice guidelines and reflects further on how guidelines developers can plan for effective implementation. This chapter also considers actions of local organizations in adapting national guidelines, problems of conflicting or inconsistent guidelines, and efforts to translate guidelines into medical review criteria.

In **Chapter 8**, the committee presents its views on the strengths and weaknesses of current efforts to develop and use clinical practice guidelines. It offers a research agenda and, more generally, proposes a framework for future development, use, evaluation, and improvement of clinical practice guidelines. A particular focus is strategies to assess the soundness of existing and future guidelines.

Several appendices help illustrate or elaborate on points raised in the

text. [Appendix A](#) provides 16 diverse examples of guidelines and related materials. [Appendix B](#) presents a provisional instrument for assessing practice guidelines, a document that was reviewed independently of the report according to NRC procedures. [Appendix C](#) presents rosters of the committees and panels involved with or contributing to this study.

SUMMARY

The recent surge of interest in clinical practice guidelines was born of frustration about seemingly uncontrollable increases in health care expenditures combined with grave doubts about the real value of that increased spending. Very high expectations for what guidelines might do to control costs and improve the value or quality of care are, however, giving way to a more pragmatic appreciation of the potential and limitations of guidelines.

The challenge to this committee was to provide a constructive analysis of current efforts to develop, use, and evaluate guidelines and to propose a framework for the future that offers realistic potential for improving the caliber and effectiveness of these efforts. This chapter has provided definitions and a context for the description, analysis, and recommendations that follow.

About this PDF file: This new digital representation of the original work has been recomposed from XML files created from the original paper book, not from the original digital typesetting files. Page breaks are true to the original; line lengths, word breaks, heading styles, and other typesetting-specific formatting, however, cannot be retained in the PDF file. Some typographic errors may have been accidentally inserted. Please use the print version of this publication as the authoritative version for attribution.