INTRODUCTION — Clinical practice guidelines are recommendations for clinicians about the care of patients with specific conditions. They should be based upon the best available research evidence and practice experience.

The Institute of Medicine defines clinical practice guidelines as "...statements that include recommendations, intended to optimize patient care, that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options" [1].

Based on this definition, guidelines have two parts:

- The foundation is a systematic review of the research evidence bearing on a clinical question, focused on the strength of the evidence on which clinical decision-making for that condition is based.

- A set of recommendations, involving both the evidence and value judgments regarding benefits and harms of alternative care options, addressing how patients with that condition should be managed, everything else being equal.

As an example, the US Preventive Services Task Force's recommendations for colorectal cancer screening [2] were published with two separate background papers: a systematic review with 129 references [3] and a decision analysis [4].

Formal advice on how clinicians should manage patients is not new. Local opinion leaders or the authors of widely-circulated review articles have had a powerful influence on practice decisions. What has changed about modern clinical practice guidelines is that they are based on systematic reviews of the evidence, are often endorsed by national organizations, undergo intensive review, and are circulated across international boundaries and specialties.

Guidelines have largely focused on the effectiveness of interventions. More recently, they have paid more attention to the size of the effect and the balance between effects on the one hand and harms and costs on the other, as well as on the feasibility of following guidelines. Another emerging development is the concept of individualized guidelines, whereby risk factors specific to the individual patient, rather than population-based risk factors, are incorporated into tools weighing risks and benefits to guide treatment decisions [5]. While such individualized guidelines have the potential to improve quality of care and lower healthcare costs, limitations to their practice application at this time include the availability of patient-specific data, validated disease models and risk calculators, and the potential impact on workflow. Guidelines developed in the United Kingdom by the National Institute for Health and Care Excellence (NICE) involve patients and caregivers in the development process and explicitly include patient choice and cost-effectiveness as factors in determining recommendations [6].

USE AND USERS — Guidelines are intended for clinicians, to help them take better care of patients. However, others involved in the health care system also use them, for better or worse. Insurers and administrators may use guidelines to set policies on quality and payment for care, by translating guideline recommendations into performance measures that are then used to assess the delivery of care. As an example, "pay for performance" (also called "pay for quality") programs are being developed to link physician payment to quality of care, as measured by guideline-derived parameters of clinical care. They are an effort to deal with rising costs and the need for quality improvement. The concept is popular, but there is not a strong body of research evidence to confirm its promise [7].
Lawyers may use well-accepted national guidelines in malpractice litigation, arguing that physicians who have not followed such guidelines without a good reason are negligent. Malpractice claims of this sort are not commonplace so far, in part because of a multiplicity of guidelines for the same condition, which tend to differ from each other, so that none can be taken as a community standard. Similar concerns about use of performance measures as evidence for medical malpractice litigation have also not been realized, since such measures report aggregate rather than patient-specific data and do not meet evidence standards [8].

Guidelines are suggestions for care, not rules. There will always be individual patients who should be managed differently. Legitimate reasons for this include biologic differences in drug metabolism, immune response, or genetic endowment; the presence of comorbid conditions; available resources determined by the social and economic environment of medicine at the local level; and patient preferences [9]. However, most patients do fit the recommendations in guidelines, and this should be reflected in practice patterns.

**RECOGNIZING CREDIBLE GUIDELINES** — Thousands of guidelines have been published. They vary in quality; some are masterpieces of clinical scholarship and judgment while others are naïve or self-serving. The most trustworthy guidelines can be recognized by the best practices for the guideline development identified by the Institute of Medicine (IOM) [10]:

- Has an explicit description of development and funding processes that is publicly accessible
- Follows a transparent process that minimizes bias, distortion, and conflicts of interest
- Is developed by a multidisciplinary panel comprising clinicians; methodological experts; and representatives, including a patient or consumer, of populations expected to be affected by the guideline
- Uses rigorous systematic evidence review and considers quality, quantity, and consistency of the aggregate of available evidence
- Summarizes evidence (and evidentiary gaps) about potential benefits and harms relevant to each recommendation
- Explains the parts that values, opinion, theory, and clinical experience play in deriving recommendations
- Provides a rating of the level of confidence in the evidence underpinning each recommendation and a rating of the strength of each recommendation
- Undergoes extensive external review that includes an open period for public comment
- Has a mechanism for revision when new evidence becomes available

Many guideline developers are revising their processes on the basis of the IOM recommendations. The National Guideline Clearinghouse (NGC) in the US has evaluated the IOM report and conducted a pilot study examining guideline developer's awareness and perceptions of the IOM standards and their intentions to meet them.

While the IOM has noted that all of these criteria should be met in order for a guideline to be judged trustworthy, measuring the "trustworthiness" of guidelines, based on how closely they adhere to these or other standards, has been proposed [11]. Additionally, questions have been raised about the feasibility of implementing all of the IOM's criteria in settings where resources may be constrained. An international network, the **Guidelines International Network (G-I-N)**, representing 51 countries and 104 organizations in 51 countries, is working to develop consensus regarding minimal standards for guideline development [12,13].

Although specific aspects of how best to implement guidelines development differ between IOM and G-I-N, generally the two bodies agree on the basic elements essential to develop high-quality guidelines, including:

- Utilizing a systematic literature review
- Establishing transparency and disclosing the methods used for all development steps
- A multidisciplinary development group
These standards have been both applauded and criticized, with the biggest issue related to the complexity and practicality of implementing the proposed standards.

**Expertise** — Guidelines prepared by panels representing the full range of expertise bearing on the clinical question are more likely to avoid the biases and blind spots of members who are all from a given specialty [14]. These panels should represent a wide range of expertise and may include, in addition to primary and subspeciality physicians, representatives from allied health sciences, public health specialists, decision analysts, economists, consumers, and ethicists. Expertise in interpreting research evidence is a necessary component of the panel. It can be contributed by members who also have other expertise, or a methods expert can be included.

The US Preventive Services Task Force makes special effort to balance the interests of its guidelines panels and its guidelines are carefully prepared [15]. In the United Kingdom, evidence-based guidelines developed by the National Clinical Guideline Centre (NCGC) for implementation in the National Health Service are also rigorously developed over a period of about 2.5 years from inception to release and incorporate decisions based upon clinical and cost-effectiveness [16]. Guidelines prepared by specialty and subspecialty societies usually involve a narrower spectrum of interest limited to members in the society [17].

**Incorporating patient perspectives** — Since the focus of clinical guidelines is on patient outcomes, and the importance of achieving or avoiding specific outcomes may differ between patients, including the perspectives of patients in guideline development is important [18]. Almost all interventions influence multiple positive and negative outcomes. Trade-offs need to be made between specific benefits and specific harms: the impact of an intervention on improving positive outcomes and generating negative outcomes (eg, side effects, adverse events, costs).

For some healthcare decisions, the values most patients place on these positive and negative outcomes can be inferred with confidence, such as the mortality benefit from the use of beta blockers following a myocardial infarction, compared to the side effects (eg, lethargy and sexual dysfunction) of beta blockers. These decisions often involve recommendations for which there is strong evidence of benefit for the intervention.

However, many healthcare decisions involve positive and negative outcomes for which the relative importance is not so clear cut, and for which the evidence of benefit is weaker. In such circumstances, such as the decision to treat lower urinary tract symptoms (LUTS) in men with a transurethral resection of the prostate or the decision to screen for breast cancer in women aged 40 to 49 years, patient values related to the importance of different outcomes are more likely to vary. Individual patients may differ in their priorities to diminish bothersome urinary symptoms or avoid sexual dysfunction, or their desire to minimize risk of a delayed cancer diagnosis versus the desire to avoid biopsy. Priorities can vary sufficiently between two patients with identical clinical circumstances such that each might best maximize their individual health outcomes by choosing exactly opposite courses of care.

Guideline developers should consider patient preferences in formulating recommendations and explicitly identify assumptions made regarding patient preferences for various outcomes. For some conditions, there is strong literature regarding patient values for competing outcomes (such as prolonged life versus sexual function in the treatment of prostate cancer). For many other conditions, the literature related to patient values is less robust and there is no agreement on how best to identify patient preference. Directly soliciting patient input as part of the guideline development process may be a reasonable solution.

**Evidence-based** — Credible guidelines are based on a systematic review of published research that is likely to include reports of all scientifically credible studies that bear on the question at hand. The systematic review should
follow explicit ground rules for identifying all relevant studies and judging their scientific strength and should make transparent the rationale for their decisions [19]. Expert opinion and usual practice not supported by research evidence may be included, but should be labeled as such and not take precedence over stronger evidence [20]. The process of developing guidelines can be used as an impetus to encourage clinical research when evidence is not available [21,22].

High-quality evidence is often not available. In a review of guidelines from the Infectious Diseases Society of American (IDSA), for example, only 14 percent of the 4000 recommendations were supported by the highest-level evidence [23]. A study of American Heart Association/American College of Cardiology guidelines found that nearly half of recommendations were based on the lowest level of evidence [24].

Even when data from well-designed and well-implemented randomized controlled trials are available, they often are not directly applicable to the populations, interventions, or outcomes addressed by the guideline being developed [25].

The aim of practice guidelines to optimize patient management may be even more relevant in circumstances where evidence is equivocal and the clinician is less certain about the choice of strategy. In such circumstances, the synthesis of carefully weighed opinions of experts may be particularly helpful, considered in the context of individual patient factors and adapted as needed. Transparency is essential, so that it is clear to readers what the quality or strength is of the evidence supporting the guideline recommendation.

Grading guidelines — Guidelines should provide an assessment of the strength of each individual recommendation. A common approach is to grade the strength of the evidence and the strength of the recommendation separately.

The Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) system [26] is gaining increasing acceptance [27-31]. In GRADE, grades have two components, a two-level representation of the strength of recommendation (strong or weak), and a four-level representation of the quality of the evidence (high, moderate, low, and very low).

- **Strength of the recommendation:** A recommendation is a strong recommendation to do (or not do) something, where the benefits clearly outweigh the risks (or vice versa) for nearly all patients. UpToDate uses a number 1 to reflect a strong recommendation. A weak recommendation is made either when risks and benefits are more closely balanced or are more uncertain. UpToDate uses a number 2 to reflect a weak recommendation.

- **Quality of evidence:** Assessment of evidence quality in GRADE reflects confidence in the estimates of benefits, harms, and burdens. GRADE can be implemented with either four levels of evidence quality, or with three levels such that the "low" and "very low" categories are combined. UpToDate uses three levels and uses a letter (A, B, or C) for high, moderate, or low/very low quality evidence. High quality of evidence typically comes from well-performed randomized controlled trials or other overwhelming evidence (such as well-executed observational studies with very large effects). Moderate-quality evidence typically comes from randomized trials with important limitations, or from other study designs with special strength. Low-quality evidence typically comes from observational studies, or from controlled trials with very serious limitations. Very low-quality evidence typically comes from nonsystematic observations, biologic reasoning, or observational studies with serious limitations. The UpToDate implementation of GRADE is shown in the table (table 1).

The US Preventive Services Task Force (USPSTF) has used a rating system that is uniform across the multiple conditions it reviews. The system features separate ratings for the strength of recommendations and the quality of evidence (table 2 and table 3). (See "Preventive care in adults: Recommendations").

Consider outcomes and implementation — Guidelines should take into account not only whether an effect of an intervention exists beyond chance but also other clinically-relevant factors, such as:
Clinical practice guidelines

- The magnitude of effect
- Harms from the intervention
- Convenience and side effects
- The clinical skills necessary to carry out the intervention successfully
- Patient preferences
- Cost
- Cost-effectiveness
- The work force necessary to implement the recommendations

The role of and potential variation in patient preferences in particular is now recognized as being of great importance for many, if not most, healthcare decisions, and guidelines on most topics should explicitly state what are the assumptions about patient preferences that were used in reaching recommendations.

Recency — Most guidelines are produced only once or, as with the USPSTF or National Cholesterol Education Program (NCEP), are updated at several-year intervals [32]. In fast-moving fields, such as treatment of HIV infection, these guidelines may be out of date in less than a year.

In a study of guidelines sponsored by the US Agency for Healthcare Research and Quality, more than three-fourths of guidelines needed updating [33]. A study of 100 quantitative systematic reviews found that new findings with impact on the review outcome occurred within two years of publication for 23 percent of the reviews [34]. The median time for "survival" of an analysis was 5.5 years. Thus, it is reasonable to assume that any guideline that has not undergone review with updating as needed within five years of publication, in the absence of strong justification, should not be accepted as current. The National Guideline Clearinghouse requires evidence that a guideline has been developed, reviewed, or revised within five years for inclusion of the guideline in their listing [35].

Sponsoring society — Statements formally endorsed by respected national bodies are subject to scrutiny by health professionals and sometimes by the public through coverage in the popular media. These organizations have a strong incentive to safeguard their reputation by having their guidelines stand up to scrutiny. However, the imprimatur of a sponsoring organization does not necessarily guarantee quality.

Review — The best guidelines undergo careful review by representatives of the sponsoring organization other than panel members. Better yet, experts outside the organization are asked to review and endorse the guidelines.

Conflict of interest — The guideline should report conflicts of interest (COI), financial or others, bearing on the guideline for each member of the panel [14].

The Institute of Medicine Committee on Standards for Developing Trustworthy Clinical Practice Guidelines recommends written disclosure of any commercial, noncommercial, intellectual, institutional, patient or public activity pertinent to the guideline scope [1]. It also recognizes that for some guidelines, a degree of COI might be unavoidable in panel participants (such as relevant clinical specialists whose income is related to providing services pertinent to the guideline) but that these members should be a minority of the panel and should not be chairs or co-chairs. In addition, the Guidelines International Network recommends that COI should be publicly disclosed, updated regularly, and no one with relevant COIs should decide the direction or strength of a recommendation [36].

While it is recognized that transparency is essential when participants in the guidelines preparation process have conflicts of interest, it is less clear how best to adapt the process to avoid bias related to any COI [37]. Clinicians with notable expertise in an area are both more likely to be sought out to participate in developing relevant guidelines and to participate in industry-sponsored activities such as speaker's bureaus or advisory panels. In various reports, the frequency of financial COIs among authors of clinical guidelines ranges from 35 to 87 percent [38-40].

QUALITY — Many guidelines do not meet current standards for quality (table 4). In a study of 279 guidelines by
69 different developers, the overall adherence to standards was 43.1 percent [41]. The greatest improvement was needed in the identification, evaluation, and synthesis of the scientific evidence. A study of guidelines prepared by specialty societies reported that most of the 413 guidelines studied did not meet criteria [42]. Performance improved from 1988 to 1998, but by 1998 only 5 percent met the three main criteria: description of the type of stakeholder, searches for published studies, and explicit grading of the strength of recommendations.

**Disagreement among guidelines** — Guidelines on the same clinical question by different expert groups often disagree. Usually the differences are minor; with screening guidelines, for example, they might differ on the age at which screening should begin and end or the time interval between screening examinations. Uncommonly, recommendations are very different. Two major guidelines for colorectal cancer screening were published within several months of each other [3,43]. One was by the US Preventive Services Task Force and the other by a consortium of the American Cancer Society, the US Multisociety Task Force, and the American College of Radiology. The two guidelines included different sets of screening test options (three options for one, seven for the other) and preferences (the ACS guidelines stated that tests that detected adenomas were preferred over those detecting cancers only). (See "Screening for colorectal cancer: Strategies in patients at average risk", section on 'Guidelines for average-risk screening'.)

Some guidelines have triggered acrimonious public criticisms. Examples are guidelines by the US Preventive Services Task Force that diverged from some professional society guidelines and did not recommend mammography screening for women 40 to 49 years of age [44] and that recommended against PSA screening for prostate cancer [45]. A review of eight guidelines on screening asymptomatic patients for peripheral arterial disease found conflicting recommendations from different organizations, with differing interpretations of the evidence base that was mostly related to the role of testing in symptomatic patients, but not to asymptomatic people [46].

Disagreement may be a barrier to acceptance of guidelines. However, in one study of this question, the extent to which clinicians agreed with each other regarding intervals for cancer screening was not related to the extent to which guidelines agreed on recommended screening intervals [47].

Disagreement among recommendations is not necessarily a sign of poor quality. A weak evidence base may lead to various conclusions. Guidelines may disagree because of the value system of the panel that developed them. One study showed that surgeons tended to favor more aggressive cancer screening intervals than family physicians and internists and that gynecologists favored more aggressive screening for cancers occurring in women [47]. The best possible care of patients may also vary according to availability of resources, health priorities, and social environment. As an example, screening guidelines tend to be less aggressive in Canada and the United Kingdom than they are in the United States.

Guidelines can differ for bad reasons too. Politics have intruded on some guidelines processes, notably those for breast cancer in women under age 50 years. Some panel members or sponsoring societies have a strong financial conflict of interest. Some panels do not include members with expertise in the interpretation of research evidence.

**Pluralism** — The United States has opted for guidelines prepared by a variety of competing organizations, most in the private sector, rather than a single version prepared by the federal government. While the multiplicity of guidelines, and their differences, can be confusing, the absence of a single standard for practice allows for innovation and protects physicians from being held to a single community standard in the absence of strong evidence on the best practice.

Other countries have a single organization developing guidelines. Such organizations include the National Institute for Health and Care Excellence (NICE) in the UK, the Institute for Quality and Efficiency in Health Care in Germany, and the Dutch College of General Practitioners in the Netherlands.

**FINDING GUIDELINES** — The National Guideline Clearinghouse (NGC) is a public resource maintained by the Agency for Healthcare Research and Quality (AHRQ). The NGC provides structured, standardized summaries of guidelines from most major medical organizations and clinical specialty societies. To be included in the NGC...
compendium, guidelines are required to meet specific criteria (eg, incorporate systematic review and an assessment of benefits and harms) based on the Institute of Medicine (IOM) definition of a clinical practice guideline.

Guidelines can also be found at the Web sites of the sponsoring organizations (table 5), and by entering "clinical practice guidelines" into search engines such as savvysearch.com, metacrawler.com, scholar.google.com/, or northernlight.com. "Practice guideline" can also be set as a limit term for type of article when using PubMed to search the US National Library of Medicine. None of these databases sorts the guidelines by quality.

OTHER TYPES OF PRACTICE GUIDANCE

Appropriate use criteria — Appropriate use criteria (AUC, sometimes referred to as "appropriateness criteria") are a variation on clinical practice guidelines, but differ in several ways [48]:

- Appropriateness is tailored as much as possible to the specific characteristics of individual patients.
- AUC cover a broader array of specific conditions, sometimes hundreds for a given test or treatment decision, to encompass the majority of practice situations. Appropriateness may relate to individual patient demographic characteristics, clinical history, risk scores, and/or symptoms and signs.
- AUC are based upon scientific evidence where possible but, in part as a result of the breadth of specific clinical conditions addressed, rely more on expert opinion than do rigorous clinical practice guidelines.
- Panels that formulate AUC are typically comprised of representatives from the specialties that manage patients with the clinical question under consideration.
- A common approach is to rate specific clinical scenarios (or potential indications for an intervention) on a 1 to 9 scale by each panelist before and after a group face-to-face discussion, and the median of the panel ratings is then used to designate each clinical scenario as appropriate, uncertain, or inappropriate.
- AUC are intended not only to help clinicians make sound clinical decisions, but also to educate patients and improve the effectiveness, efficiency, and equity of care.

Implementation of some AUC has been facilitated by the development of calculators or tables, in which the clinician can input specific patient information. As with guidelines, AUC are not intended to replace clinical judgment.

The AUC approach has been used primarily by subspecialty societies concerned with cardiovascular, rheumatologic, and pulmonary conditions, and associated imaging procedures. AUC have also been developed for several surgical procedures [49]. AUC are included in the National Guidelines Clearing House.

Guidance statements — The American College of Physicians (ACP) has issued several "guidance statements" that are based on reviews of other guidelines and not a de novo systematic review of evidence. These address conditions such as breast cancer screening in women aged 40 to 49, prostate cancer screening, or colorectal cancer screening for which there is general agreement on all or most of the eligible trials, but difference in how the evidence from those trials is interpreted. The ACP reviews existing guidelines on the topic and appraises them using criteria from the "Appraisal of Guidelines, Research, and Evaluation (AGREE II)" [50]. The ACP then formulates its own guidance, based on a review of the literature, recommendations in existing high-quality guidelines, and considering the needs and values of its general internist membership and their patients.

ATTITUDES AND ACCEPTANCE — In general, clinicians are most likely to accept recommendations from their own specialty society, less likely to trust those prepared by government agencies, and least likely to believe in guidelines prepared by managed care organizations and insurance companies. In one study of internists’ attitudes toward guidelines, 82 percent had confidence in guidelines prepared by the American College of Physicians and only 6 percent in those prepared by Blue Cross and Blue Shield [51]. Attitudes related to who prepared guidelines appear to be independent of the scientific validity of the guidelines themselves, although this has not been
Clinicians disagree on whether clinical practice guidelines promote "cook-book medicine" [52] or evidence-based medicine. In one study of a national sample of internists, the majority believed that guidelines are good educational tools, a convenient source of advice, and are intended to improve the quality of care and decrease health care costs [51]. However, a majority also believed that guidelines are biased, oversimplified, and rigid, likely to decrease physician reimbursement, challenge physician autonomy, and decrease physician satisfaction.

Physicians are more likely to trust guidelines if they have had a hand in developing them. However, it is neither practical nor necessary for local physicians to build new guidelines from the ground up if guidelines by national groups already exist. Much of the work needed to assemble the evidence, which can take thousands of hours, has been done. Local physician groups should critically review guidelines prepared by national groups, even if they are well-done, and consider modifying them to suit their local situation. The conditions of practice, resources, and patient preferences may be distinctive enough from practice to practice to lead to somewhat different recommendations. Also, local modifications of guidelines foster local buy-in.

**Potential benefits/problems of guidelines** — Evidence-based, carefully developed and updated guidelines provide many potential benefits:

- Synthesis of the literature by experts
- Clear recommendations for translating the evidence base into clinical application to foster best practice
- Opportunity to evaluate the outcomes of implementation in the "real world" setting

However, several aspects of guidelines and their implementation need to be recognized as potential problems:

- The challenge of keeping guidelines updated when the literature changes.
- The potential for inappropriate use of guidelines as tools for legal or restrictive administrative purposes.
- Difficulty accessing guidelines at the point of care: many are lengthy, or specific components relevant to a patient are not readily searchable or retrievable.
- Lack of coordination between guideline development groups, generating differing recommendations.
- Potential for conflict of interests.
- Application of guidelines developed to address a specific condition to patients with multiple morbidities [53,54].

**EFFECTS OF GUIDELINES ON PRACTICE** — Experts in guidelines and evidence-based medicine urge clinicians to use them as a guide, rather than as a set of rules or cookbook, and to tailor clinical decisions to individual patients. Clinical practice guidelines generally are not intended as quality measures. Guidelines often do not adequately account for severity of illness, patient preferences, or clinical judgment to be able to be used as quality measures [55].

Simply providing guidelines seems to improve practice, but the effects are small [56]. It is clear that there are significant gaps between guidelines’ recommendations and practice. As an example, in one study of practice in the United States, it was estimated that 68,000 deaths could be prevented if six recommendations from guidelines for the management of heart failure were closely followed [57]. There are many reasons why clinicians do not adhere to guidelines’ recommendations. In one report using focus groups of general practitioners in the Netherlands to explore reasons why recommendations from 12 national guidelines were not followed, the most frequently perceived barriers to adoption were identified as disagreement with the recommendations, environmental factors based on organizational constraints, lack of knowledge of the recommendations, and unclear or ambiguous guidelines [58]. In other studies, recommendations were more likely to be followed when they were supported by clear evidence, were compatible with existing norms and values, did not require new skills or change in practice...
routine, were less controversial, and were stated in specific actionable terms [59,60].

Thus, the "actionability" of a guideline is an important attribute. While well-formulated guidelines can be an invaluable tool to guide best practice in medicine, they should not alone be considered a complete plan for quality improvement. Rather, they need to be delivered in the context of a program to engage patients and clinicians in appropriate decision-making, supported by implementation strategies involving systems enhancements, clinical reminders, other quality improvement and decision support tools, and outcomes measurement and feedback.

However, even when such support is provided, it remains challenging to change clinician practice. In a randomized trial related to guidelines for the management of nonvariceal upper gastrointestinal bleeding, 43 hospitals were randomly assigned to an intervention (physicians received published consensus guidelines, algorithms, written reminders, and participated in multidisciplinary guideline education groups and case-based workshops) or control (received guidelines and algorithms) [61]. At one year, guideline adherence was not significantly different between the intervention and control hospitals, with adherence below 10 percent in both groups.

**Implementing practice guidelines** — Increasing attention has been focused on how best to disseminate guidelines and foster adoption of their recommendations, once developed [62,63]. A framework for the study of guideline implementation("implementation science") includes five principal domains: characteristics of the intervention, the outer setting, the inner setting, characteristics of the involved individuals, and the process of implementation [64].

Strategies to facilitate implementation of practice guidelines have been proposed and include [65]:

- Guidelines should incorporate a checklist of prioritized specific interventions
- Identify barriers to adoption, and design supports to address specific barriers
- Integrate guidelines for common coexistent conditions
- Identify systems and technological solutions to promote adherence with recommendations
- Develop transdisciplinary teams (clinical epidemiology, implementation science, systems engineering) to study ways to foster best practices.

The Institute of Medicine (IOM) recommends that guidelines that are based upon strong evidence should be worded so that it is possible to evaluate whether care followed recommendations [66]. The IOM suggests that guidelines be structured in format, vocabulary, and content to foster use of computer-aided decision supports by guidelines users.

**SUMMARY AND RECOMMENDATIONS**

- Clinical practice guidelines are recommendations for clinicians about the care of patients with specific conditions. Guideline development should involve a systematic review of the research evidence related to decision-making for the targeted condition/question, and recommendations about patient management based on the evidence and value judgments that should be explicitly identified as such. (See 'Introduction' above.)

- Guidelines are suggestions for care, not rules. There will always be individual patients who should be managed differently for reasons, including biologic differences (in drug metabolism, immune response, or genetic endowment), comorbidities, availability of resources and cultural differences, and patient preferences. For some healthcare decisions, differences in patient preferences for various health outcomes can mean that there is no one course of care that can be strongly recommended. (See 'Use and users' above.)

- Guidelines vary widely in quality. Credible guidelines involve: development by a panel representing a full range of expertise; an unbiased systematic review of the evidence; grading the strength of the evidence and recommendation; incorporation of multiple relevant factors including feasibility, harms, costs, and patient preferences; and a process for ongoing review and updating (table 4). (See 'Recognizing credible guidelines' above.)

- While well-formulated guidelines can be an invaluable tool to guide best practice in medicine, they should not
alone be considered a complete plan for quality improvement. Guidelines need to be delivered in the context of a program to engage patients and clinicians in appropriate decision-making, supported by implementation strategies involving systems enhancements, clinical reminders, other quality improvement and decision support tools, and outcomes measurement and feedback. Practice guidelines are not performance measures. While performance measures are often developed from practice guidelines, the latter usually specify in advance what would be optimal care, while the former assess retrospectively whether care met some minimum threshold of performance. (See 'Effects of guidelines on practice' above.)

ACKNOWLEDGMENT — The editorial staff at UpToDate would like to acknowledge Robert H Fletcher, MD, MSc, who contributed to an earlier version of this topic review.

Use of UpToDate is subject to the Subscription and License Agreement.

REFERENCES


17. Brook RH. Practice guidelines: to be or not to be. Lancet 1996; 348:1005.


## GRADE for practice guidelines

<table>
<thead>
<tr>
<th>Grade of recommendation*</th>
<th>Clarity of risk/benefit</th>
<th>Quality of supporting evidence</th>
<th>Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>Benefits clearly outweigh risk and burdens, or vice versa</td>
<td>Consistent evidence from well performed randomized, controlled trials or overwhelming evidence of some other form. Further research is unlikely to change our confidence in the estimate of benefit and risk.</td>
<td>Strong recommendation, can apply to most patients in most circumstances without reservation</td>
</tr>
<tr>
<td>1B</td>
<td>Benefits clearly outweigh risk and burdens, or vice versa</td>
<td>Evidence from randomized, controlled trials with important limitations (inconsistent results, methodologic flaws, indirect or imprecise), or very strong evidence of some other form. Further research (if performed) is likely to have an impact on our confidence in the estimate of benefit and risk and may change the estimate.</td>
<td>Strong recommendation, likely to apply to most patients</td>
</tr>
<tr>
<td>1C</td>
<td>Benefits appear to outweigh risk and burdens, or vice versa</td>
<td>Evidence from observational studies, unsystematic clinical experience, or from randomized, controlled trials with serious flaws. Any estimate of effect is uncertain.</td>
<td>Relatively strong recommendation; might change when higher quality evidence becomes available</td>
</tr>
<tr>
<td>2A</td>
<td>Benefits closely balanced with risks and burdens</td>
<td>Consistent evidence from well performed randomized,</td>
<td>Weak recommendation, best action may differ</td>
</tr>
<tr>
<td>Quality of Evidence</td>
<td>Recommendation</td>
<td>Evidence</td>
<td>Confidence in Estimate</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------</td>
<td>----------</td>
<td>------------------------</td>
</tr>
<tr>
<td>High quality evidence</td>
<td>controlled trials or overwhelming evidence of some other form. Further research is unlikely to change our confidence in the estimate of benefit and risk.</td>
<td>depending on circumstances or patients or societal values</td>
<td></td>
</tr>
<tr>
<td>2B Weak recommendation Moderate quality evidence</td>
<td>Benefits closely balanced with risks and burdens, some uncertainty in the estimates of benefits, risks and burdens</td>
<td>Evidence from randomized, controlled trials with important limitations (inconsistent results, methodologic flaws, indirect or imprecise), or very strong evidence of some other form. Further research (if performed) is likely to have an impact on our confidence in the estimate of benefit and risk and may change the estimate.</td>
<td>Weak recommendation, alternative approaches likely to be better for some patients under some circumstances</td>
</tr>
<tr>
<td>2C Weak recommendation Low quality evidence</td>
<td>Uncertainty in the estimates of benefits, risks, and burdens; benefits may be closely balanced with risks and burdens</td>
<td>Evidence from observational studies, unsystematic clinical experience, or from randomized, controlled trials with serious flaws. Any estimate of effect is uncertain.</td>
<td>Very weak recommendation; other alternatives may be equally reasonable</td>
</tr>
</tbody>
</table>

* GRADE can be implemented with either three or four levels of quality of evidence. UpToDate implements three levels and uses numbers and letters to represent strength of recommendation and quality of evidence respectively.

Graphic 90325 Version 2.0
# USPSTF grading scheme for recommendations

The U.S. Preventive Services Task Force (USPSTF) grades its recommendations according to one of five classifications (A, B, C, D, I) reflecting the strength of evidence and magnitude of net benefit (benefits minus harms).

<table>
<thead>
<tr>
<th>Classification</th>
<th>Recommendation</th>
<th>Evidence and Balance of Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A</strong></td>
<td>The USPSTF strongly recommends that clinicians provide [the service] to eligible patients. The USPSTF found good evidence that [the service] improves important health outcomes and concludes that benefits substantially outweigh harms.</td>
<td></td>
</tr>
<tr>
<td><strong>B</strong></td>
<td>The USPSTF recommends that clinicians provide [this service] to eligible patients. The USPSTF found at least fair evidence that [the service] improves important health outcomes and concludes that benefits outweigh harms.</td>
<td></td>
</tr>
<tr>
<td><strong>C</strong></td>
<td>The USPSTF makes no recommendation for or against routine provision of [the service]. The USPSTF found at least fair evidence that [the service] can improve health outcomes but concludes that the balance of benefits and harms is too close to justify a general recommendation.</td>
<td></td>
</tr>
<tr>
<td><strong>D</strong></td>
<td>The USPSTF recommends against routinely providing [the service] to asymptomatic patients. The USPSTF found at least fair evidence that [the service] is ineffective or that harms outweigh benefits.</td>
<td></td>
</tr>
<tr>
<td><strong>I</strong></td>
<td>The USPSTF concludes that the evidence is insufficient to recommend for or against routinely providing [the service]. Evidence that the [service] is effective is lacking, of poor quality, or conflicting and the balance of benefits and harms cannot be determined.</td>
<td></td>
</tr>
</tbody>
</table>


Graphic 75920 Version 1.0
# USPSTF grading scheme for quality of evidence

<table>
<thead>
<tr>
<th>Good:</th>
<th>Evidence includes consistent results from well-designed, well-conducted studies in representative populations that directly assess effects on health outcomes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fair:</td>
<td>Evidence is sufficient to determine effects on health outcomes, but the strength of the evidence is limited by the number, quality, or consistency of the individual studies, generalizability to routine practice, or indirect nature of the evidence on health outcomes.</td>
</tr>
<tr>
<td>Poor:</td>
<td>Evidence is insufficient to assess the effects on health outcomes because of limited number or power of studies, important flaws in their design or conduct, gaps in the chain of evidence, or lack of information on important health outcomes.</td>
</tr>
</tbody>
</table>

The USPSTF grades the quality of the overall evidence for a service on a 3-point scale (good, fair, poor).


Graphic 73780 Version 1.0
# Criteria for trustworthy clinical practice guidelines

<table>
<thead>
<tr>
<th>Standard</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Transparency</td>
<td>Guidelines should include an explicit description of process and funding.</td>
</tr>
<tr>
<td>2. Conflict of interest</td>
<td>Conflicts of interest for the guidelines development group should be managed by reporting, exclusion, and divestments.</td>
</tr>
<tr>
<td>3. Members of the guidelines development group</td>
<td>The group should be multidisciplinary and balanced.</td>
</tr>
<tr>
<td>4. Review of the literature</td>
<td>The guideline should be based on systematic reviews of the literature.</td>
</tr>
<tr>
<td>5. Rating strength of evidence and recommendations</td>
<td>Each recommendation should be accompanied by the underlying reasoning, potential benefits and harms, the evidence and its quality, the contribution of values and experience, rating of the level of confidence in the evidence and the strength of the recommendation, and differences of opinion regarding recommendations.</td>
</tr>
<tr>
<td>6. Presentation of recommendations</td>
<td>The guideline should state precisely the recommended actions, when they should be performed, and how they could be measured for evaluation of compliance.</td>
</tr>
<tr>
<td>7. External review</td>
<td>The guidelines should be reviewed by the full spectrum of relevant stakeholders. The general public should have an opportunity to review the guidelines before they are final.</td>
</tr>
<tr>
<td>8. Updating</td>
<td>Guidelines should state date of publication and evidence review and be updated when new, clinically-important evidence is available.</td>
</tr>
</tbody>
</table>


Graphic 68090 Version 3.0
### Web sites for major guidelines

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Centers for Disease Control Guidelines for Sexually Transmitted Disease</td>
<td><a href="http://www.cdc.gov/std/treatment/">http://www.cdc.gov/std/treatment/</a></td>
</tr>
<tr>
<td>American College of Physicians/American Society of Internal Medicine</td>
<td><a href="http://www.acponline.org">http://www.acponline.org</a></td>
</tr>
<tr>
<td>American Cancer Society</td>
<td><a href="http://cancer.org">http://cancer.org</a></td>
</tr>
</tbody>
</table>

Graphic 80337 Version 1.0
Disclosures


Contributor disclosures are reviewed for conflicts of interest by the editorial group. When found, these are addressed by vetting through a multi-level review process, and through requirements for references to be provided to support the content. Appropriately referenced content is required of all authors and must conform to UpToDate standards of evidence.

Conflict of interest policy