Updating Clinical Practice Guidelines

By

Joseph A. Scattoloni

A Master's Paper submitted to the faculty of
The University of North Carolina at Chapel Hill
In partial fulfillment of the requirements for
The degree of Master of Public Health in
The Public Health Leadership Program.

Chapel Hill
2002
ABSTRACT

Context: Clinical Practice Guidelines (CPGs) are an expanding area of healthcare, but may be worthless and in fact hazardous if they are not current. Virtually no published methods or specific guides to updating them exist, and there have been few published tracts that document the different originating organizations' actual methods concerning guideline review and revision.

Objective: To document, compare, and evaluate the CPG development and update methodologies of seven professional physician organizations.

Participants: Six physician organizations were selected from the National Guideline Clearinghouse (NGC). An additional physician organization's methods are introduced from the author's experience working for said organization.

Methods: A preliminary questionnaire was sent to selected organization representatives (six). Data was collected through informal, semi-structured phone interviews. A single organization (USPSTF) was sampled through the author's one-month practicum experience developing and applying review methods to guidelines possibly requiring update.

Results: All organizations were very concerned with the rapidly changing healthcare environment and have goals of keeping their guidelines up to date. However, these goals generally have not been translated into explicit operational methods. One of seven have specified update interval; three of seven have systematic processes for updating; four of seven have systematic processes for doing an update. None have explicit methods for selecting topic experts. Time estimates for updates range from twelve to thirty-six months (eighteen month average).

Conclusion: Guidelines have been promoted as a necessary means of improving healthcare for both individual patients and for the system at large, but the methods in place for systematically reviewing and updating previously created guidelines are sub-optimal. Attention to this problem is increasing as efforts to develop new methods have started appearing both in the literature and in several organizations (herein discussed). These labors in addition to increasing information technologies should allow ongoing review and updating of guidelines in the future.
INTRODUCTION

Over the past two decades, evidence-based clinical practice guidelines (CPGs) have developed from an idea to a reality, and now play an ever-greater role in the practice of medicine. The concept of CPGs has been in place for many years as specialty organizations tried to codify general practice parameters to standardize and strengthen patient care practice. However, these guidelines were almost always arrived at through "expert" opinion stemming from panel discussions. In fact, they might be organized and written in a single session or day. This approach, informal consensus development\(^1\) or global subjective judgement\(^2\), was the precursor process for what became known as "guidelines" during the post-WWII years. These guidelines were not evidence-based, but rather reflected simple expert opinion in assessing whether a practice was appropriate. They also lacked methods regarding evidence assessment how consensus was achieved. A more formal consensus development method was advanced and practiced during the seventies and early eighties by several groups (the NIH, AMA, and Harvard Community Health Plan).\(^1\) The most formal of these methods (instituted by the RAND Corporation) still used panels or surveys of experts, but the process was more systematized and documented, often including actual grading scales for topics.\(^3\) It nevertheless failed to expressly connect recommendations with evidence.

At the same time, another methodology was being developed which directly correlated guideline recommendations with demonstrable underlying evidence. The necessity of incorporating clear scientific evidence into guideline
development was acknowledged as early as 1980 by the American College of Physicians, who initiated the Clinical Efficacy Assessment Project (CEAP) to develop guidelines based on the assessment of clinical literature by topic experts. Similar programs began to appear in specialty organizations and insurance groups. Even more stringent or explicit connections between evidence and recommendations were required in the methodology espoused by the Canadian Task Force on the Periodic Health Examination in 1979. Following suit, the U.S. Preventive Services Task Force established similar rules of evidence in 1984, which have become more explicit and systematic since. In these latter systems, a coded schema rates the quality of evidence.

Ultimately, clinical practice guidelines were thought to be a means of improving the quality of patient care. They could improve health outcomes by improving the consistency of care, by disseminating information to patients and doctors alike, and by influencing public policy. By the late eighties and early nineties, however, there was rapidly growing further interest in the development of these guidelines. Health care costs were becoming a greater concern for all parties -- payers, practitioners, and patients -- and guidelines were seen as a means of lowering costs and improving efficiency by reducing inappropriate care and controlling geographic variation in practice characteristics. With the need to curb costs fueling a political firestorm at the federal level in the late eighties, practice guideline legislation was adopted and mandated as an alternative to lowering costs by way of expressed expenditure targets. It was felt that practice guidelines could couple curtailing inappropriate or redundant care with improving
general health practices. The Agency for Health Care Policy and Research (AHCPR) was created under HCFA, and the government became a major player in the development and dissemination of CPGs.

For all these reasons, the number of guidelines began to expand dramatically through the 1990s. MEDLINE searches demonstrate the rapid increases in published CPGs and in articles generally concerned with CPGs.

![Number of Published Guidelines per Year (1981-2001)](image)

This from a MEDLINE search crossing the exploded MESH headings of Family Practice, Pediatrics, Psychiatry, Nursing, all surgical specialties, and all medical specialties, limiting by individual years to English, Human, and being either a Practice Guideline or a Guideline.

![Published Articles with "Practice Guideline" as Subject](image)

This from a MEDLINE search using the MESH heading Practice Guidelines and limiting by English language, Human, and per year.
Updating Clinical Practice Guidelines

With this increase in the number of guidelines, there has been a dramatic rise in the number of organizations who create and disseminate such practice parameters. Creators include any number of different allied health professional societies (from the Academy of Ambulatory Foot and Ankle Surgery to the American Academy of Pediatrics to the Society of Nuclear Medicine), government agencies, insurance agencies and coalitions, patient advocacy groups, and others. The actual use of these guidelines is hard to assess due to their limited access, as well as the difficulty trying to compare and contrast guidelines on similar topics.

To provide a common pathway to evidence-based guidelines, AHCPR utilized the "new" Internet technologies to produce a central site -- the National Guideline Clearinghouse (NGC) -- that would gather guidelines from disparate sources and provide detailed and comparative methodological information on these guidelines. The NGC became operational in January of 1999 under the Agency for Healthcare Research and Quality (AHRQ) (formerly AHCPR). The U.S. National Guideline Clearinghouse currently lists 116 contributors and over 900 guidelines. Some of these groups publish more than thirty guidelines. Others have contributed a single guideline. Browsing the NGC (at www.guideline.gov) assures some minimum standards for the guidelines revealed, but the actual number of circulating guidelines beyond the NGC is in the several thousands.

So, it begs the question, are all these guidelines valid and of good quality? When the federal government entered into the guideline business, it contracted the Institute of Medicine to advance some general guideline concepts to inform the
new venture. The Institute of Medicine (IOM) standardized the definition of guidelines as "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances. The key word here is "systematically," for it lies at the root of the historic shift from an unplanned, almost anonymous process to the highly structured, evidence-based methods today expected.

By their scientific nature, laboratory experiments and clinical trials strive to be inherently neutral processes that generate outcome or result data. It is the interpretation of this data that propels medicine forward and impacts patient care. The more quality evidence or data collected, the easier it is to make judgements. Evidence reports, tables, or balance sheets organize empirical data; systematic reviews document, collate, and summarize this data; guidelines interpret this data, weighing the strength of evidence for harms and benefits and costs. Each of these steps amplifies the possibility for error as it relies more on human judgement and less on quantitative assessment.

David Eddy, one of the originators of the theory behind evidence-based CPGs, states that CPGs or "policies" should be accurate, accountable, predictable, defensible, and usable. They should veer as far from subjectivity as possible. Explicit formal analysis techniques are the best protection against subjectivity, inaccuracy and bias. The closer the guideline comes to meeting all these objectives, the higher its quality. Yet each of these objectives relies on different aspects of the method used to generate the guideline. This expands the resources required to assure and increase each guideline's quality.
The generally accepted steps in guideline development are: identifying and refining the guideline topic; selecting and running a guideline development group; identifying and assessing the evidence; translating this evidence into a CPG; disseminating and eventually updating the CPG. Whenever possible, each of these steps should conform to a rigorous method that can be explicated and reproduced. The systematic review is the heart of the CPG's development; it shapes the guiding questions and then identifies and analyzes the evidence from which recommendations are then made. It is probably the most systematized aspect of guideline development for which there exist formal directions or illustrations. Other aspects of the process are not clear and are still being advanced. For instance, criteria for delineating both composition and selection of a topic group's membership can be inadequate or haphazard. In the absence of clear evidence and concerning certain aspects of quantifiable evidence (e.g. generalizability), expert opinion or clinical consensus will be used to make recommendations. However, as Shekelle et al. state, "There is currently no optimal method for...deciding how to collect and assess expert opinion." Another incapacitating aspect of the process is that at present there exist only "passive" methods for disseminating and implementing guidelines, which result in minor or negligible changes in professional behavior. Finally, until quite recently there have been no systematic attempts to formulate a method for reviewing and updating guidelines.

Clinical practice guidelines then can vary widely in their construction. Furthermore, there are the considerable costs associated with CPG development.
The more "systematic" the process, the greater the resources and time required. The costs of publishing or producing these CPGs generally do not include the extensive efforts of the large number of persons contributing clinical, methodological, epidemiological, and writing expertise.\textsuperscript{10,14} A single high quality systematic review contracted by AHRQ from its national Evidence Practice Centers (EPCs) costs approximately $250,000.\textsuperscript{16} The demands of strict methodology and elevated cost limit the quality of many CPGs, but do not necessarily curb their creation. Many guidelines are promulgated to meet the narrow needs of the guideline developer and remain rooted in the tradition of consensus or opinion. In fact, many North American healthcare organizations "purchase commercially produced guidelines" that emphasize practice behavior that generates resource savings/profit.\textsuperscript{6} The waters have been further muddied by creation of "practice advisories" in which recommendations are created where insufficient evidence may exist. These advisories are usually arrived at via the older, traditional methods of clinical consensus, but have gained greater prominence due to their close "approximation" to evidence-based guidelines.

CPGs should not be viewed as a panacea to the various ills of a stressed healthcare system. There are intrinsic limitations regarding the absence or poor quality of scientific evidence and the still evolving methods for gathering and interpreting the evidence that exists. There are also limitations stemming from external pressures being brought into the guideline development process. Various contrasting agendas may "guide" the focus or process beyond patients' needs to exigencies to control costs, serve societal needs, or protect special interests.\textsuperscript{7,17}
This admixture of inadequate method and vested interest may diminish or even thwart the guideline's supposed primary objective (per the IOM) of improving the quality of individual patient care. Flawed CPGs may harm individual patient care in any number of ways: by overlooking good or using inaccurate evidence; by being inflexible; by being outdated and maintaining obsolete practices/treatments; by dismissing patient preferences; by disrupting doctor-patient relationships; or by inappropriately directing public policy and thus limiting patient education or resources.6

As demonstrated above, the number of guidelines is rapidly increasing. The varying forces that drive this proliferation of guidelines seem to be feeding from their own energy or inertia. If one or two well-constructed guidelines are good, then many guidelines of unproven quality are even better. For all the focus on evidence-based practice informing these guideline constructs, there have been few attempts to scientifically assess the results of these efforts. Treatment and diagnostic modalities are critically appraised through highly structured trials, but CPGs, which to a greater and greater degree generate patient care plans, are only beginning to be subjected to such scrutiny. So far the results are mixed. A systematic review by Grimshaw and Russell considered fifty-nine evaluative studies and concluded that "explicit guidelines do improve clinical practice," but that the guidelines' performances varied considerably.18 A more recent review was concerned only with CPGs' effect on primary care outcomes. Thirteen of ninety-one identified studies met the inclusion criteria (reported patient outcomes); the
results indicated that implemented CPGs "in primary care settings [do] not consistently result in improvements in clinical outcomes."19

It should be clear that CPGs have become a fixture in health care for the immediate future. While there appears to be great potential for such guidelines to improve patient care quality and clinical effectiveness, the development process and eventual implementation are still evolving and have not yet been truly evaluated. Many concerns need to be addressed. Some of these are being contemplated and investigated to a greater degree than others. My recent experience working with one of AHRQ's EPCs centered my attention on a guideline issue that has been generally overlooked in the literature until just recently -- the updating of guidelines, or when and how to proceed with this review and revision.

In the lengthy chain of steps assuring a CPG's ultimate benefit, updating is a final step that cements a circular process. Review and update are ultimately crucial to the defined mission of improving individual patient care and acknowledging the evolving secondary concerns regarding consolidating or controlling the costs/resources associated with this care. Medicine evolves more and more rapidly. Over two million articles, and more than 17,000 biomedical books are published annually.20 The basis for guidelines, even if originally superbly researched and constructed, shifts and changes. Without amending and clearly announcing these amendments to CPGs, outmoded practices and technologies will be perpetuated, patient care will be compromised, and costs will increase due to resultant inefficiency and variations in practice. Furthermore, all
the resources that went into the original guideline will have been wasted. There is no point in creating a CPG unless one plans to maintain and update it.

During the 1990s, resources were spent on lengthy development practices for CPGs. Methodologies were constructed and refined for creating the best possible guidelines with the available information or evidence. The necessity for constant or frequent review of the changing literature was acknowledged in theory, and this acknowledgment led to the general specification that every guideline should include a scheduled review date. But for all of the last decade's interest in and work on guideline development, it is the absence of any considered update methodology that is notable. A scheduled or binding review date can result in wasted resources or inappropriate care depending on the velocity of change in the concerned field of practice. Questions then remain surrounding how to decide when an update is appropriate, and then to potentially gauge the type of update required. Must a full repeat of the initial development process take place, or might the guideline only need itemized updating?

Paul Shekelle and associates recently noted that no one had previously addressed how new information should trigger the update process. They approached the issue by first suggesting an outline of basic principles that might indicate the need for a CPG update (see box); then they proposed a model for

- Changes in evidence on the existing benefits and harms of interventions.
- Changes in outcomes considered important
- Changes in available interventions
- Changes in evidence that current practice is optimal
- Changes in values placed on outcomes
- Changes in resources available for health care

guideline assessment (including decision algorithm). The focus of this assessment was on identifying new evidence and assessing whether this new evidence warranted updating a given guideline. Two means were advanced to identify the new evidence: consulting expert opinion and performing focused literature searches. Experts would include previous guideline developers as well as other topic experts and generalists with guideline expertise. The literature review would be limited (in order to save time and resources) to reviews, editorials, commentaries, and any new guidelines on the given topic. Shekelle's group theorized that any new evidence significant enough to annul an existing national CPG must be known to established topic experts or have been published in significant journals with accompanying editorial comment due to its nature as a "sentinel marker" of new evidence. The assemblage of information may then inform the decision about whether or not an update is appropriate. No specified process was formulated to guide this decision. Shekelle's group noted that it is inherently subjective and relies on the clinical and methodological expertise of the deciding group.

Shekelle and his associates later implemented their proposal using as test subjects seventeen of the original CPGs established by the AHCPR in the early and mid-nineties. They completed and examined a strict process that measured the identification of new evidence that might compel a major, minor, or no update, and calculated survival analysis of the rate at which guidelines became outdated. Evidence was gathered per the previous discussed methods and evaluated by the paper's authors. When a guideline's principal diagnostic or therapeutic
recommendations were called into question by the new evidence, a "major" update or guideline withdrawal was warranted. If only secondary recommendations within the guideline were challenged and needed changing or refinement, a "minor" update was warranted. Otherwise, the guideline remained valid without changes. In this study, seven guidelines required major updates, six needed minor updates, three were still valid, and one could not be conclusively decided. One half of the guidelines became obsolete in 5.8 years (CI of 5.0-6.6), and 3.6 years (CI of 2.6-4.6) was the point at which 90% of guidelines were still valid.

This was a seminal study and established a new way of considering and carrying out the review and revision of guidelines. However, there are still many troublesome aspects. Even in this regimented study, only rudimentary methods for expert selection were advanced. No distinction in the update process has been established between a "major" and "minor" update; indeed, no definition of an actual method to perform the updates has been advanced. This theorized process presupposes well-constructed guidelines where key questions informing the process are easily or already established. An analytic framework and resultant key questions are essential to gathering pertinent evidence, and this proposition and study fail to provide detailed steps regarding the literature search process. New modalities or ways of thinking about the topic may require a new "framework," and difficulties and time increase with the need to create analytic frameworks.
USPSTF Update Process

As a medical and public health student, I participated at a national Evidence Practice Center (EPC) that assessed whether aging guidelines from the U.S. Preventive Services Task Force (USPSTF) Guide to Clinical Preventive Services (2nd Ed.) required revision. Facing the NGC's requirement that all guidelines be validated at least every five years, the USPSTF sought a means of reconsidering older recommendations to determine their status as either still current or out of date and needing to be revised. Working with AHRQ, the contracted Research Triangle Institute/University of North Carolina (RTI/UNC) EPC was charged with initiating a process for just such an evaluation. A work group comprised of EPC staff and librarian, preventive medicine residents, and a medical student examined sixteen previously un-reviewed chapter topics over a four-month period (mid-September 2001 to mid-January 2002). Except for the work group leader, these participants were unfamiliar with CPG development and review.

The project proceedings may be summarized as follows:

- No clear or central organization
- Empirically propelled experimental process
- No clear initial oversight by the USPSTF
- BMJ article by Shekelle et al. was a catalyst but not a blueprint for process
- Participants worked individually on topics
- Proposed process (while not necessarily followed for assigned 16 topics) was documented for the USPSTF (see box 1)
Observations concerning the assigned task and its process were assembled from my own experience writing five topic assessments, participating in weekly or bimonthly work group meetings, and informally surveying other work group participants.

**Proposed Updating Process for USPSTF (12/15/01)**

1) Read old chapter, draw Analytic Framework, define "critical" key questions (KQs)
2) Send old chapter to 2-3 experts in the field, asking standard questions about the existence of new evidence, ongoing studies, and their opinion of whether revision is needed
3) Do Cochrane search for systematic reviews and new RCTs for critical KQs
4) Do MEDLINE search for systematic reviews, editorials, commentaries, and new evidence (since 1994) on critical KQs, using restrictive search terms and inclusion criteria designed to find high quality evidence
5) Search the NGC and key organizations for new guidelines at variance with the USPSTF
6) Examine editorials, commentaries, and reviews for new high quality evidence not found elsewhere
7) Examine references from experts and other guidelines
8) Write brief abstract of high quality new evidence for each critical KQ
9) Briefly summarize new evidence for each critical KQ, including information from experts and other guidelines
10) Task Force subcommittee examines summary of evidence and reviews key studies if needed
11) Present summary of new evidence (1-2 pages) to entire Task Force at a meeting for discussion; topic is assigned a priority code for revision update

**SIGNIFICANT ASPECTS of PROCESS**

**Defining Critical Questions**

Guidelines are created to assist in answering questions. To reassess them demands a clear understanding of each guideline's context and the questions it is attempting to acknowledge. This skeletal or analytic framework (AF) becomes the basis for strategies developed to create and, in our case, re-assess guidelines through the critical questions they attempt to answer. Several difficulties were noted:

- Overarching or key questions were difficult to agree upon.
- Uncertainty hindered the ability to decide where to draw the line when incorporating topic material for assessment:
• Certain topics were extremely broad.
• Some topics had sustained rapid changes underlying concepts/application.
• The introduction to and application of analytic framework (AF) came too late in the process to make a difference.
• Working alone, each participant felt uneasy about the reliability of the process.

Literature Searches

Our supposed task was to gather in the most efficient or economical manner the best quality evidence to inform a recommendation. It begs the question: how is "high quality" evidence defined and then identified? The definitions and hence the type of evidence may vary per topic. "Counseling outcomes" may be much more difficult to quantify than the diagnostic effects of a blood test in "screening outcomes." The Shekelle et al. criteria enlist editorials and commentaries in "major journals" to route this search directly to "quality" evidence. Our work group defined "quality" by randomized controlled trials (RCTs) or other closely controlled trials and our searches encompassed "all" the literature. Two real problems stood out:

• Whether to limit or expand searches. Tied to the aforementioned difficulties with developing topic frameworks and key questions, searches were initially quite limited. However, with evolving expansion of many topics' crucial points, additional searches crossing separate headings for treatments, or harms, or other "buried" information became necessary.
• When to halt searches. This raises the question of whether the undertaken process is simply to document a need for updating, or to actually perform a significant portion of the update itself. If the literature provides one or
two clear examples of high quality evidence changing the supposed guideline, isn't the process completed?

**Experts**

Experts may redefine a topic's parameters, identify new evidence, and assess the evidence in the context of such parameters. Two to four experts were selected for each topic and were mailed (by post and/or e-mail) copies of the previous Task Force guideline and specific questions for the applicable topic/chapter (see box 2).

The response rate was poor (see box 3) and variation of expert response was extreme from three page references and full text unpublished articles to single sentence replies. No means was implemented to check if method of contact or format of questions would have changed the outcome. It should be noted that no compensation was offered to these experts and this may have increased the response rate.

There was no established method for selecting or validating these experts. Networking or word of mouth advised the selection of some, while the majority of

---

**Excerpted from letter sent to Experts**

We will be grateful if you could advise us concerning the following questions. Do you have knowledge of:

- New evidence (i.e., since 1994) or developments in the field that relate to the prior USPSTF evidence review and recommendation;
- New evidence that is expected to be published in the near future;
- Any changes in available interventions;
- Any changes in the evidence on the benefits and harms of treatment.

In addition, any guidance or references to materials published in the past seven years that you believe will illuminate this particular clinical topic will be appreciated.

Finally, in your opinion, how great is the need to update this chapter in the next 1 to 2 years - great, some but not urgent, or not really needed?

Box 2

**Consulted Experts' Response**

16/48 (33%) = Total Response Rate
4/18 (22%) = Response rate for experts outside the U.S.
4/16 (25%) = Response rate to Email
1 mo - Average time to respond for those who responded

Box 3
experts resulted from scanning article authors from literature searches. Work force members used their own methods to select experts. This increased the inconsistencies and probable bias, and may have contributed to the poor response rate.

Oversight by issuing or contracting agency

The work group was ultimately contracted by the USPSTF to serve their requirements. This actually created operational difficulties within our work group:

- There were no specific instructions regarding the scope of each chapter or guideline review.
- When there was interaction, the rules kept changing, causing continual process modifications.

In fact several designs would have greatly facilitated the process:

- The contracting agency (USPSTF) should provide direction and expectations at the very beginning of the process.
- This agency should at outset provide immediate input into the development of AF, key questions, and search parameters.
- Guidelines ought to document original or subsequent AFs and key question evolution.

Results

The sixteen topic briefs or summaries were presented to the USPSTF in late January 2002, and their disposition may be viewed in box 4. Three of the sixteen, while close to resolution, required some further work. While under discussion, there was no definition advanced for the specific types of changes or required further work entailed in tweaks, minor, or major updates. Rough
estimates among work force members regarding time demands (literature searches, article pulling and review, meetings/conferences, and writing) were:

- Approximately fifteen hours for simple topics (e.g. screening for testicular or pancreatic cancer).
- More than forty to fifty hours for complicated or evidence rich topics (e.g. tuberculosis, AAA, or screening ultrasonography in pregnancy).

**Comments**

It is true that we did not implement a systematic trial of a specified process like the process described in Shekelle et al's JAMA article; in fact, it was an evolving process. Nevertheless, it raises some important issues. First, are key questions and analytic framework of existing guidelines easily identified? They are requisite for the construction of an effective and evolving search for new information. Indeed, they go to the heart of understanding the basis of the original guideline and how it must evolve with the changing medical literature. To streamline the process and avoid a vacillating definition of review goals, early definitions must be established. Moreover, there should be agreement among independent reviewers concerning this framework and the resulting key questions.
Studies should be undertaken to evaluate this process for inter-observer agreement.

The foundation for any necessary changes to a guideline derives from a reliable assertion of new evidence. Therefore, studies should also be implemented to discern the inter-observer agreement (reliability) between reviewers who use these "established key questions" to identify credible or high quality new evidence.

For all the gathered evidence and all procedures used to minimize the introduction of bias to the process, decisions about how to develop criteria or identify and assess evidence are inherently subjective. "Experts" on the specific topics and on guideline methods will be required, but a specified selection process should be formulated. Experts should not be picked through a haphazard manner, nor should they be selected through "word of mouth" in an insular network. Once again, if good criteria are created for approaching this selection, then there should be inter-observer agreement among reviewers.

Incorporating both literature review and expert opinion expands the ability to capture and evaluate relevant new evidence, but it should be considered whether there is a significant difference between the results of expert query versus focused literature search. What is the best manner of contacting and questioning selected experts? Was our poor response rate due to the method of contact? And how many expert opinions are enough to sufficiently inform the process?
The UPDATE PROCESS at other ORGANIZATIONS

Given my experience in the EPC work group, I wondered how or even if other organizations that generate CPGs were addressing the "update question." I resolved to begin a small-scale evaluation of this question.

METHODS

To more fully understand the processes surrounding the development and updating of clinical practice guidelines, I interviewed representatives from a sample of six organizations that write guidelines. These professional medical societies or organizations had at least ten guidelines published on the National Guideline Clearinghouse. I assumed that physician organizations would incorporate the most stringent, well-developed, and reproducible methodologies or processes regarding guideline construction and maintenance. Selections after these baseline criteria were not methodical but ordered according to leading or earliest position in the NGC and a probable bias regarding organization reputation or influence. The six organizations were:

- The American Academy of Pediatrics (AAP)
- The American Association of Clinical Endocrinologists (AACE)
- The American College of Cardiology (ACC)
- The American College of Physicians/American Society of Internal Medicine (ACP-ASIM)
- The American Psychiatric Association (APA)
- The American Society of Anesthesiologists (ASA).
The contact information for each of the organization's headquarters was located using Internet web searches. I telephoned the central operator at each organization, introduced myself, and asked for the e-mail and direct extension to the most appropriate person to speak with regarding the organization's maintenance of clinical practice guidelines. Before every interview, the assigned representative received by e-mail a standard template of queries (shown here) to better inform them of the general parameters of my questioning. The subsequent conversations were informal and semi-structured around my list of questions. In several instances, the calls were conducted in two sessions due to respondent time constraints. What follows are general notes and impressions from these conversations.

### RESULTS

(See Table 1 for overview of organizations)

<table>
<thead>
<tr>
<th>GUIDELINE UPDATE QUESTIONNAIRE for Creating Organizations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. When were guidelines last updated? (mo/yr) (never updated)</td>
</tr>
<tr>
<td>2. When will the next update occur? (mo/yr) (not yet determined)</td>
</tr>
<tr>
<td>3. Is there a schedule dictating when specific guidelines are to be updated?</td>
</tr>
<tr>
<td>4. If there is not a schedule, is there a specified system your agency uses to review guideline topics in order to check whether they need an update?</td>
</tr>
<tr>
<td>• Who or what group conducts this review to decide whether an update is warranted? (intra-agency) (outside agency)</td>
</tr>
<tr>
<td>• Does this same person or group then undertake the update process itself?</td>
</tr>
<tr>
<td>5. If an update is required, does your agency utilize a standard methodology to direct the update process?</td>
</tr>
<tr>
<td>• Is the process written?</td>
</tr>
<tr>
<td>• Can you send it to me?</td>
</tr>
<tr>
<td>6. Please describe briefly this methodology if possible — the following questions may or may not inform your description.</td>
</tr>
<tr>
<td>• Are topics updated through a literature search?</td>
</tr>
<tr>
<td>• Is a complete systematic review undertaken?</td>
</tr>
<tr>
<td>• What are the parameters or criteria of such a search? (single key quest) (multiple key quests) (only RCTS, Reviews, Edits/Comments) (some combo of article or study types)</td>
</tr>
<tr>
<td>• Are topic updates informed through the solicitation of expert opinion?</td>
</tr>
<tr>
<td>• How are experts selected?</td>
</tr>
<tr>
<td>• How many are selected?</td>
</tr>
<tr>
<td>• How many generally respond?</td>
</tr>
<tr>
<td>• After researching the topic, is there a system describing or categorizing the amount of revision each guideline requires? Who does this revision? Who must approve it? (complete update via systematic review) (complete update per collected evidence or expert opinion) (moderate update) (minor update) (no update required) (other categories)</td>
</tr>
<tr>
<td>7. How long does it take to update a guideline?</td>
</tr>
<tr>
<td>8. Can you estimate the costs for updating a topic? ($) (person hours)</td>
</tr>
<tr>
<td>9. Who makes the final decision whether a guideline update is adopted and what it will recommend? (committee) (individual) (outside agency)</td>
</tr>
</tbody>
</table>
At the AAP, I spoke with Carla Herrerias, MPH, Senior Health Policy Analyst, who told me that the first guidelines from the AAP began appearing in 1994. At that time, it was established that guidelines would be assessed every three years and be either "revised, retired, or re-affirmed." Such assessment would incorporate the same methodology as the initial development of the said guideline with the only concession being an amendment of the employed literature search strategy by limiting the dates of the search to a minimal overlap with the previous search parameters. Thus, each guideline update is a complete "re-development" of the original evidence-based guideline. Every guideline must begin such reassessment every three years. All the AAP's guidelines from the years 1994-1999 have been through or are being put through such a process.

Ms. Herrerias stated that evidence-based Practice Guidelines take approximately twenty-four to thirty-six months to develop and publish. This time estimate holds for the "updates" or three-year re-assessments. She emphasized that each guideline was different depending on its complexity and the new evidence amassed, as well as the dynamics of each guideline workgroup. However, literature review and analysis generally takes twelve months, and an additional eighteen to twenty-four months are required for discussion, writing, and adoption of the new guideline (or its reaffirmation or retirement). Because of the extensive and time-consuming method, the process of reassessment usually begins soon after the guideline is published (within twelve to eighteen months).
Guideline topics originate in the National Quality Improvement Committee, and workgroup members (who are considered experts for said topic) are selected via nominations from the AAP’s national committees and/or sections. These nominations are reviewed by the Quality Improvement Committee and finally approved by the AAP board of directors. This sub-committee or workgroup usually numbers ten to fifteen persons including an epidemiologist and, unless conflicts arise, these members comprise the sub-committee assigned to continual updating. Members are asked to continue "observing" the literature in preparation for updating the guideline. They are replaced individually as needed through the nomination process mentioned above.

A comprehensive methodology example and flow chart of the actual process from beginning to end was sent to me by Ms. Herrerias and is shown in appendix 1. The literature review is described as "in-depth" and is based on the criteria specified in the guideline. From Ms. Herrerias' description, it is a full systematic review of the literature although we did not discuss specific types of studies considered applicable to these reviews. The experts queried are those on the sub-committee itself. As alluded to above, a three tier system is used to categorize an update -- revision (essentially a complete rewrite), re-affirmation (keeping the old guideline intact), and retirement (withdrawing the entire guideline from further sanctioned use). All these changes or recommendations originate in the sub-committee and go to national specialty committees and societies and other outside agencies for review. Final changes are made by the subcommittee and are reviewed first by the Steering Committee on Quality
Improvement and Management, then the Board of Directors, and finally the AAP Executive Committee.

Ms. Herrerias mentioned that costs for updating guidelines are extremely hard to estimate and that there have not been any attempts to quantify these costs up to now. The AAP relies on the volunteer activities of many clinicians inside and outside the AAP for panel participation and assistance with each literature review. There are paid epidemiologists, librarians, and support staff involved. When further queried, she estimated a minimum of $30,000 for each update, but the number is probably much higher. The AAP has been partnering with or contracting the services of AHRQ and its EPCs to a greater and greater degree for guideline development and review purposes, and she noted that their estimates for a complete systematic review in the process of guideline development was somewhere around $250,000 per guideline.

Asked if she thought the present system was effective, Ms. Herrerias said that it was quite thorough and, in fact, perhaps too thorough. A great deal of time and resources (monetary and clinician time) are expended on the update process, much of which may be redundant. She is not sure that each guideline needs to be completely revamped, and there has been increasing concern within the AAP about the resources being utilized. Discussion has ensued concerning other more streamlined ways to conduct the update process -- using a more focused literature search, polling experts about new evidence rather than performing a literature search, reviewing and/or changing sections of guidelines instead of the whole. However, none of these changes are presently on the table to be implemented. In
her capacity, Ms. Herreries interacts with other organizations that develop and update guidelines, and she stated that the difficulties and questions concerning updating methods are presently mounting and gaining greater attention.

AMERICAN ASSOCIATION of CLINICAL ENDOCRINOLOGISTS

Sissy Crabtree is the AACE's Director of Communications. She stated that the AACE has published thirteen guidelines over the past eight years. There is no specific schedule or process for reviewing these guidelines in order to assess the need for update. It is an informal process with a suggested "scheduled review" or "expiration date" of five years from publishing date. In fact, there is not a specified methodology for the creation of these guidelines; there are simply approximate standards for their construction.

Guidelines originate when the AACE Board of Directors selects a chairperson, who is a topic expert (and sometimes a Board member), to oversee a work group of experts chosen from candidates nominated by this chairperson and the Board. No detailed selection method exists for these individuals. This work group usually comprises some five to ten members. The specific methodology the work group follows is explicated by the chairman and does not necessarily hold true for other work groups. Literature search parameters may vary, as may assessment techniques per the personal views of this chairperson.

Ms. Crabtree did send me the "AACE Standards on Medical Guidelines for Clinical Practice," which was developed from information from the AMA (appendix 2). This description seems to serve a purely suggestive function and does not elaborate or explicate any portion of the process. For instance, #10 on
this list deals specifically with update concerns by proposing that an expiration date or scheduled review date (within five years) be written into the guideline. Furthermore, it states, "[A] system is established to monitor the emergence of [new] information, which may necessitate revision." It does not propose such a system, and in fact, as described above, the AACE does not follow a prescribed system either, granting each chairperson their own autonomy in judging new evidence toward the decision to update older guidelines. Ms. Crabtree describes the processes actually in place as quite "informal." The "standards" document also states that guidelines should outline methods for evaluating their own impact and using this evaluation in the revision discussion and implementation. There is nothing in place at the AACE to even begin such an evaluation.

When guidelines (or updates) have been completed in the work group, they go to the Publications Committee, which checks for technical accuracy and content, and the chairperson of this committee has a dialogue with the work group chairperson to make further revisions. From there it goes to the Medical Editor of the journal who reviews it before publishing. All members of the AACE Board (33-35 members) have the "opportunity" to review it and comment, but this is not a presumed or explicit part of the process.

Guideline development at the AACE takes around twelve months, and Ms. Crabtree said that updates take similar amounts of time but may be quicker depending on the chairperson. As revealed at other agencies, costs are extremely difficult to estimate, but she said that publishing costs (including medical writer) of each guideline run around $25,000. Of course, there are costs associated with
support staff and the many volunteer hours of work group members' time. Ms. Crabtree suspected that each guideline's total cost might hover around $100,000.

AMERICAN COLLEGE of CARDIOLOGY / AMERICAN HEART ASSOCIATION

Charlene May, Director of Document Development and Practice Guidelines, explained that she is involved with the document management "process" and that I should also speak with Paula Thompson, MPH (chief methodologist/analyst). Ms. Thompson is involved with "methods" and has been instrumental in the construction of the ACC's web site documenting their guideline development methods. I spent time conversing with both Ms. May and Ms. Thompson.

The ACC published its first guideline in 1981. At present there are seventeen that the ACC/AHA have produced, but the process of development for these CPGs has evolved over this time period. Ms. Thompson reported that it has only been in the last several years that a research analyst or methodologist has become a mandatory component in each of the ACC guideline-writing committees. Prior to this change, these committees policed themselves regarding their "undocumented" methods. This new emphasis has allowed for a clearer documentation of the methodologies employed by the ACC/AHA in the guideline development/update process. This documentation is available to peruse on the organization's website, www.acc.org (downloaded to appendix 3). In fact, it is quite thorough in its description of the methods employed for guideline development and updates.
Originally, guidelines were to be reviewed for update every five years; then the interval changed to two years, but this has been amended to one year after publication and every subsequent year thereafter. The system to determine whether an update is required begins with two persons, the writing committee's chair and its research analyst. These two identify ten to twelve important journals relevant to the topic (advised by the initial guideline's literature analysis) and then monitor them over the course or at the end of the year following initial publication. Articles that might change a guideline recommendation or signal a major shift in thought process would indicate the need for an update. The entire writing committee is also surveyed every year for their input regarding any momentous changes in the evidence that would signal the need to update. The writing committee's formal recommendation goes to the parent committee (Task Force on Practice Guidelines), which decides whether to update or not.

When required, the actual update process proceeds down one of two paths, either "update" or "revision." Either path incorporates the same strict methodology of evidence procurement and analysis as the initial guideline development. "Updates" stem from a necessity to change one or more of the recommendations of a specific guideline. These changes are incorporated into the original document with clear reference to the new evidence. They are published in three forms: the original or preceding guideline with old sections struck out and the new in bold print (this demonstrates the changes); the new altered version; and a summary of how and why the old version has been changed.
"Revisions" are complete rewrites of the guideline and are used when the document no longer reflects the standard of care it supposedly covers. Ms. Thompson cited the recent "Heart failure" revision (2001), which required a re-conceptualization of the guideline's purpose due to fundamental changes in the understanding of the condition. This guideline was started from scratch; new clinical objectives were debated and developed, and these raised the questions that guided a full systematic review. But such total revisions are fewer now as increased staff and more constant review of the literature allow for flexible and continued quality "updating" of guidelines. Still, Ms Davis explained, only two "updates" are allowed each guideline before a total revision is mandated.

The Task Force on Practice Parameters creates writing committees composed of eight to twelve members by first determining guideline topics and then (with the assistance of the ACC and AHA presidents) nominating and selecting a chairman who will assist the Task Force in the selection of the remaining committee members from the Colleges. Up to three "outside" specialty organizations (not necessarily physicians) are invited to send a working representative to the writing committee. These are the only members who submit CVs and interview. The committee is responsible for the guideline's content following the set methods. As with other organizations, there is no specified or objective method for the selection of these members. They are the "acknowledged tops" in their academic or clinical endeavors. The writing committee stays intact in order to facilitate continued review of the literature and the possibility of updates, but membership changes each year after the first year's anniversary with
Updating Clinical Practice Guidelines

Joseph A. Scattoloni

revolving reappointment of one-third and two-thirds of committee members in opposite years.

Each guideline or update making its way out of the Task Force undergoes a lengthy peer-review process. First there is "content" review by multiple (as many as ten) "experts" selected by each of the writing committee members. Their input is debated and possibly incorporated. Then there is the "official" review. The guideline draft goes to three experts from the AHA, three from the ACC, and one each from any "outside" agency included in the writing committee. The chairman is responsible for responding to all "official" reviewers and making any requisite changes with the writing committee. The AHA and ACC board of trustees (thirty members) debate this draft and hash out any further differences with the chairman and College presidents before making a resolution to accept a finished document. It should be noted that although this procedure is elaborate and quite thorough, there is nonetheless a lack of explicit criteria for the selection of these reviewers/experts.

All documents are "team-written" with each phase and every subsection divided among members so that consensus development is ongoing. Ms. Thompson admitted, however, that this also slows and potentially politicizes the process. An initial guideline development or full revision takes eighteen to twenty-four months, while updates take approximately one year or less. Only certain sections of updates demand a review and rewrite. Ms. Thompson also stated that the disparate factors involved in facilitating quick guideline reviews/updates center primarily on the interpersonal skills of the committee.
chairman, but also on group members. One member can completely logjam the work product. Finally, staff support is crucial, as a methodologist, for instance, may provide a steadiness or continuity to the group.

Ms. May stated that update costs would be impossible to estimate at this time, but that a conservative estimate of a full guideline development or revision would be around $100,000. This would include the time of one science staffer, one project manager, and some support staff, but only minimum inclusion of overhead. She reminded me that a great portion of the work is done by College members whose time is strictly voluntary and un-reimbursed except for direct expenses like travel and lodging to conference/discussion meetings. Some monies have been recently saved because the College does not publish hard copies of the full guidelines any more. It relies on electronic media for this.

AMERICAN COLLEGE of PHYSICIANS / AMERICAN SOCIETY of INTERNAL MEDICINE

The operator at the ASIM directed me to the scientific policy department where I received contact information for Vincenza Snow, M.D., Senior Medical Associate, Department of Scientific Policy. The ACP-ASIM has been in the guideline creation business since 1981 and currently supports fifteen guidelines. But its archives include forty-three previously published guidelines that have expired. Dr. Snow explained that a strict "sundowning" policy dictates that every practice guideline will be withdrawn after five years of its most recent publication unless a formal review and update results in a new publication in the Annals of Internal Medicine. So, specific guidelines do not exist for more than five years and an update is not a guaranteed event.
Dr. Snow described a loose but straightforward and constant process to review current guidelines for necessary updates. First, an abbreviated or focused literature search of the major journals (Annals, BMJ, JAMA, NEJM, and Lancet) is ongoing for anything that points to new evidence regarding the specific topic. Any such information is presented at the quarterly meetings of the Clinical Efficacy Assessment Subcommittee (CEAS) and after discussion this information may trigger a formal directive to update. These literature searches may also turn up newly published and well-regarded guidelines that may circumvent the need for any update, thus predestining the ASIM guideline to withdrawal upon expiration. Dr. Snow cited recent screening guidelines disseminated by the U.S. Preventive Services Task Force as examples of this proceeding. Finally, all guidelines are formally reviewed in committee two years before expiration through the amassed results of these focused literature searches and committee members' expertise and knowledge of the topic.

One must understand that the ASIM develops and updates guidelines in a manner slightly different from other organizations. Their guidelines originate either in an opportunistic or de novo fashion. "Opportunistic" supposes that a clear, well-supported evidence or systematic review is available or being developed relevant to a desired topic of the ASIM. Obvious examples of such reviews are Evidence-based Practice Center (EPC) reports contracted by the Agency for Healthcare Research and Quality (AHRQ). These evidence reports form the basis of "background papers" created by the original EPC authors/researchers that may then be assimilated into a policy piece written and
adopted by the ASIM. Dr. Snow used the example of "pharmacotherapy for depression." The original review of the evidence concerned patient groups beyond the Society's purvey (children, pregnant women, etc.). So, the EPC authors were asked to write a summary of their findings regarding adults.

The de novo process starts within the ASIM and utilizes members' expertise and efforts. Dr. Snow and CEAS members research (via own knowledge, literature, contacts) possible topic or content experts from within the membership who might carry out a systematic review of the evidence. The CEAS examines nominated physicians' CVs and publications and selects the task force. This usually comprises two or three people with an additional methodologist. This group is responsible for the "background paper" that will be integrated into the policy statement (or final guideline).

Whether developed via intra- or extra- organization, Dr. Snow considers the background papers through each draft. Likewise, Dr. Snow and the Scientific Policy Staff are responsible for the construction of all policy statements and the integration of these background papers. This arrangement around a single person/entity streamlines the process and avoids unnecessary procedural hitches that might develop with multiple creators. Dr. Snow acts as an advocate or gatekeeper for each of the guidelines, and through her position the ASIM guidelines acquire continuity and consistency.

Regarding updates, there is no system in place to categorize the amount of necessary revision; a total rewrite is initiated. The background or evidence authors are released after the publication of the guideline and background paper or
corresponding updates. If a further revision is required, these original authors are invited, if unwilling or unable, to recommend another group to perform the evidence assessment. Literature reviews for updates begin where the original search left off.

There is currently no written or standard ASIM methodology to direct reviewers through guideline development or update. The ASIM's history with certain groups like the EPCs informs its understanding and sanctioning of their employed methods, which include standardized scoring of the literature, evidence tables, and consensus statements where there are gaps in the evidence. Nevertheless, the CEAS does participate in and to a certain extent coaches the progress and process of the evidence analysis in the "background papers" through early and continued communication with the background or evidence authors. The CEAS endorses the topic's analytic framework, key questions, and evidence linkages, and they view and comment on each draft of these papers. Nevertheless, Dr. Snow stated that the Scientific Policy Staff is in the process of developing a "guideline of authorship" which will explicitly chart a methodology for guideline development/update. For now the ASIM's loose descriptions of its efforts and general methods (appendix 4) can be found at its website (www.acponline.org/sci-policy/guidelines).

It takes approximately eighteen months to develop these guidelines and shepherd them through the required review process. The CEAS approves the final background paper for method and content. It works with Dr. Snow to finish a policy or guideline statement that must then be approved successively by the
Education Committee and the Board of Regents. A further evaluation of the employed method and content takes place when the guideline is submitted to the *Annals of Internal Medicine*, which has autonomy in its acceptance or rejection of the College's statement.

Dr. Snow was unable to provide any estimates of guideline costs. She and her staff are working on nine to eleven guidelines in various stages at any one time, and a great portion of the work is completed through volunteerism by College members or by outside agencies (like the EPCs) working from their own budgets. She was also unable to provide any substantive information regarding the actual use or effects of the guidelines she had helped to produce. While the ASIM's Physicians' Quality Network (*Qnet*) may examine the issue of guideline dissemination, implementation, and outcomes assessment in the future, there are currently no specific plans for such studies due to lack of funding. This greatly concerns her because the time and effort going into the construction of such guidelines may be completely wasted as these practice guidelines "are left on the bookshelves," fail to be implemented, or worse, actually diminish patient care.

**AMERICAN PSYCHIATRIC ASSOCIATION**

The operator at the APA connected me with Rebecca Thaler, MPH, Senior Project Manager for Practice Guidelines. I sent her my e-mail questionnaire, and she sent me the APA's methodology guides before we finally spoke at length.

The APA's first practice guideline was completed in 1994 and eleven more have been published since then with four more new guidelines to be published in the next two years. Two previous guidelines have already undergone
Updating Clinical Practice Guidelines

Joseph A. Scattoloni

updating/revision and five more are in the midst of updating with planned publishing dates ranging from 2002 to 2004. When the APA originally published these guidelines, they had planned on mandating an update every three to five years incorporating major or minor changes depending on the shifts in or evolution of the pertinent literature/thought process. Every ten years, a complete guideline rewrite would be undertaken. Ms. Thaler explained that this projected update schedule has been remodeled. Due to the exorbitant resources required for a thorough and complete evaluation of the guidelines, the concept of making major/minor changes every three years has been abandoned for simply implementing a full revision or rewrite process when the need presses.

Who then decides when this need becomes pressing enough? Or likewise, is there a system in place to check the necessity of an update? Per my conversation, the decision or impetus to update is derived from the expertise of the Steering Committee, writing workgroup, and other clinician/researchers regarding new diagnostic or treatment modalities. There is not a system per se besides the concerns of conferring experts or associates who many times have not been formally consulted. Nevertheless, the five-year yardstick remains in place to measure approximately when serious consideration or discussion should commence concerning an update (if it has not previously been initiated).

While the update process is not "set in stone," it closely adheres to the specific methodology outlined for initial guideline development at the APA. This strategy (derived from the AMA and IOM) is outlined in the APA Practice Guideline Development Process (appendix 5) and is fairly explicit. Ms. Thaler
said that the revision or update process differs from the initial guideline development only in that literature searches may be "more abbreviated" by way of shortened search-date intervals, or specific treatment or other key questions. There are also usually fewer drafts simply because the work group has the previous guideline for reference.

From the documents provided, it is clear that the APA does try to follow a rigorous evidence-based procedure, but that "where gaps exist in the research data, evidence is derived from clinical consensus." So, expert opinion is "quantified" as part of the process through an extensive series of draft reviews by the work group, the Steering Committee, fifty reviewers "with expertise in the subject area," a variety of other APA committees/councils, and one hundred "representatives of related organizations." This scrutiny is impressive and probably more complete than most reviews, and yet there is no articulated procedure or criteria for selecting these experts.

Ms. Thaler reported that guidelines generally take more than twenty-four months to develop from scratch and approximately eighteen months to update or revise. She was unable to provide accurate costs associated with the process. There is no specific budget per guideline developed or updated, but she concedes $50,000-75,000 per update would be a very rough estimate. Of course, this does not include the number of voluntary hours by work group and other APA members, which is considerable and varies from topic to topic.

Regarding the lengthy process around guideline updates, Ms. Thaler acknowledged the difficulties implicit in staying abreast of new information in the
guidelines while adhering to the high standard required by updates incorporating the "full review" or rewrite process. This is especially difficult with guidelines like the APA's, which are quite exhaustive and run fifty to seventy pages in length. The emerging question within the APA has been how to stay current between these exhaustive reviews. A more streamlined "advisory" is being developed which would act to supplement guidelines with specific new treatment or diagnostic information derived from "limited" literature searches and expert opinion. These "Watches" would not be guidelines or carry the weight or consensus of sanctioned guidelines, but would only provide a central information hub to clinicians, and might finally assist or be part of the decision-making process informing when a full update is warranted. Such "Watches" have not been put in place but are currently being developed.

American Society of Anesthesiologists

Richard Connis, Ph.D., Chief Methodologist, Committee on Practice Parameters, discussed the ASA's procedures regarding guidelines and updates.

The ASA's first guidelines were published in 1993. Since then twelve more have been published. During the past nine years only one (sedation and analgesia) has been updated (adopted in October 2001, to be published upcoming), but three more are in the process of being updated with possible publication over the next two years. There is no specific schedule mandated for updating these guidelines, and no explicited process exists for reviewing present guidelines to assess their current applicability.
The ASA Committee on Practice Parameters was established in the early 1990s to oversee the development of all ASA guidelines. Along with other agenda items at its annual meetings, this committee decides which topics merit consideration as guideline topics and then appoints a chairperson (most often not a member of the committee) to oversee a task force numbering eight to twelve members in the development of said guideline. The committee with this chairman then nominates recognized topic experts from the United States and Canada to serve on this task force. There are no set criteria or systems for the selection of these experts. Even after guideline completion, this task force is never disbanded; it remains in place as a resource on this topic to help decide when and how to revise the guideline. One member of the task force is designated as an oversight person to communicate task force views or progress to the Parameters Committee.

Topic updating is guided by the expressed need of members of this task force or members of the Practice Parameters Committee. If committee members, through their own concern, think an update is required, it is discussed and put forward to the task force. Likewise, if task force members (the experts on the topic) feel new evidence demands assessment, they communicate with the larger Committee to start the ball rolling back to the task force. The task force makes a recommendation to the committee after meeting with all its members to discuss possible updates. So, the impetus driving an update is a two-way street between these groups, but without explicit process. Updates may occur as early as one to two years after the most recent publication. The Practice Parameters Committee must formally consider every guideline within five years of its last formal
assessments. The assigned methodologist for each task force will raise the five-year formal inquiry in the event that no one has expressed an interest in guideline revision prior to this time.

An update at the ASA entails a complete guideline revision whereby the assigned task force starts from scratch as if they were developing the initial guideline. The only difference from previous development resides in the limiting dates of the literature search, which start one year before the most recent date of publishing. The actual process or methodology of development has been previously published (appendix 6). The ASA is different from most other organizations in that it employs a three-tier process of guideline development -- directional assessment of literature syntheses, meta-analysis, and expert consensus as evidence. The initial literature search is quite broad, excluding only editorials, commentaries, correspondence, reviews, meta-analysis, and unpublished papers/presentations, and including even case reports and descriptive studies. Well-designed studies including statistical information are pooled into a formal meta-analysis. The writing process incorporates interaction between the task force writers (chairman and methodologist), task force members, and other ASA members (via internet). A long chain of open forums with input leads to a final vote by the ASA House of Delegates.

If the guideline is not adopted, then the process must start from scratch at a later date. All three sources of evidence (directional, meta-analysis, consensus) are separately considered, and all three must agree, or the guideline cannot be put forward. The inclusion of meta-analysis is quite stringent and different from most
other organizations. Dr. Connis explained that in anesthesiology outcomes are often more consistent and clear-cut than in other specialties, thereby facilitating the pooling of data. When meta-analysis cannot be completed, the resulting document will not be termed a guideline, but instead will be called a "practice advisory."

Dr. Connis said that these guidelines take two to three years to develop and are a consuming undertaking. Time and cost estimates are very difficult to assess, especially as his section is working on two to three guidelines at one time. His very rough estimates assign production costs for each guideline in the neighborhood of $150,000-200,000. There are two full-time methodologists, a part time librarian, and various support staff working on these guidelines, but a tremendous number of hours are logged by the volunteer activities of the task force and other Committee members. The task force chairman alone probably spends over two hundred hours working on his group's topic.

The ASA has no plans at this point to revamp its decision tree concerning when or how to update guidelines.

COMMENTS

While all guideline creators recognize the need to update or revise guidelines, there really is no consensus regarding how to decide when such updates are required. Of those interviewed, only the AAP mandates definite scheduled updates, and they are having great difficulties with the increasing amount of work this entails.
Furthermore, only the ACC and ACP-ASIM have expressed methods for recognizing update necessity based on changes in published evidence. Expert guidance is generally relied upon to distinguish when revisions are warranted, but this expert opinion is quite insular to the publishing organization and original work groups. I discerned no set methodology for the selection or evaluation of such experts besides general networking or acknowledgment. Of course, in relatively small specialties (e.g. the 5,000 members of the AACE) a general agreement among members as to topic expertise probably does yield the appropriate personnel. However, to my knowledge, no group has seriously evaluated whether this type of expert enlistment does indeed capture an appropriate group of individuals or whether the overall knowledge base of such a group is diminished by a subjective and potentially biased method of selection. It should be noted that this difficulty also potentially plagues the guideline's actual update process, as expert opinion is requisite for the interpretation of the evidence accumulated after any literature or systematic review.

Some form of systematic, albeit abbreviated or restricted, literature search might be used to supplement expert opinion when deciding whether new evidence recommends a guideline update. Other than the ACC and ASIM, the use of such a search has not really been considered. At the least, a study might distinguish whether adding such a literature search advises the decision to update in a different and/or more productive manner than just relying on previously called experts.
It is also clear from speaking with these persons that assigning accurate costs to guideline development or revision is virtually impossible. The estimates vary widely. This makes some sense because organizations all have different ideas and goals for their guidelines, and some topics and/or organization guidelines are more labor-intensive than others. However, cost considerations must be evaluated. Guideline revision is a continual and cyclical process, and therefore a continual drain on healthcare resources. If we are to judge whether guidelines and their required constant upkeep are a boon to healthcare, we must be able to balance their costs against what is potentially given up for them. Organizations should attempt to itemize the resources put forth in the guideline development or revision process. This might constitute another study.

For all the resources expended on CPGs, there have been only cursory attempts to quantify the degree to which this information is being disseminated and finally utilized by physicians, health agencies, and patients. There have been even fewer attempts to measure the effects or outcomes of CPGs on the practice of clinical medicine. This makes sense in that resources were initially focused on the construction and implementation of CPGs; however, the time has come to test these health care mechanisms. In my informal conversations, the lack of past efforts on this front was generally acknowledged. There is still uncertainty as to how to factually evaluate CPGs outcomes. Money is tight, so once again it is the larger organizations that are slowly undertaking the effort. The ASIM's Qnet project is setting up large scale cooperative outcome studies to evaluate guideline effects, and the ACC has recently started a division, Guidelines Applied in...
Practice (GAP), whose chief function will be "scientifically" evaluating the implementation of guidelines and whether patient care is improved.

It appeared that the larger the organization, the more thorough were its methods. The ASIM and ACC all have appreciable, if varying, efforts in place to review trends or new evidence on an ongoing basis. The AAP has scheduled updates. The ASA, while smaller, produces full-scale reviews including meta-analysis. For all its efforts, the AACE, an IM subspecialty organization, was clearly the least systematic in process and review. With only 5,000 members and limited resources, the AACE accomplishes as much as it is able; and as a physician organization, its expertise certainly bolsters the production of quality guidelines. However, these observations should serve as a caution regarding the presumed general quality of CPGs. It would be interesting to renew this sort of evaluative study between larger physicians' organizations who have produced many guidelines and smaller doctor groups that have put out only a few. Likewise, a study could be undertaken to compare physician group CPG originators with either patient advocate groups or the insurance groups that produce guidelines. Are there differences in the methods employed by these groups? Are there differences in the final products?

One of the most confusing update issues is the fine line between what constitutes an update and what constitutes the information gathering used to decide to update. The overlap between the two causes a redundancy of efforts and resources. A more fluid process of continual update may avoid this redundancy.
One possible algorithm demonstrating such a process is shown in Figure 1. A systematic development process is initiated to create the CPG. The topic writing committee or panel would remain intact or integrate a proportional rotation of members (like the ACC). Personnel trained in methods would "sift" or comb through the literature on an ongoing basis (perhaps quarterly) with clearly defined search parameters that were developed in conjunction with the initial development panel or committee. Additional annual surveys sent to systematically selected experts would supplement this information. Evidence tables or summaries could be disseminated to the topic committee each quarter and discussed at annual meetings to decide whether to leave the CPG as is, to update it by pertinent recommendation, section, or background information, or to start the process over from scratch. Additionally, a formal review for the topic panel or committee would be required five years from the most recent full rewrite; essentially, an examination of the amassed evidence from the previous five years of "sifter" literature searches and expert testimony would occur.

Problems seem to arise when long intervals stretch between evidence assessments. A greater accumulation of information complicates the appraisal of any evidence and increases the potential for redundancy between an update check and an actual update. This is especially true because a defined approach to updating has not been developed. Should an update be a re-development of the guideline, or can it be a piecemeal revision of the existing document? An ongoing information collection process incorporates updating; the information found goes straight into the update.
Even if piecemeal revision is incorporated via ongoing review, a five year assessment of the accumulated evidence is prudent to decide if the guideline should be rewritten in total. The guideline developers may lose sight of the forest by focusing on the trees. If new conceptualizations or medical practices have developed around or beyond the old topic, then the topic's foundation framework may have become irrelevant no matter how many piecemeal revisions have been made. The scheduled re-evaluation of the topic ameliorates this possibility without forcing a full-scale rewrite when it is not necessary.

The "sifters" or methodologists' collection of evidence could be conducted via searches for RCTs, or editorials, or commentaries, or expert opinion. It only needs to be established which method or combination of methods secures the best results. Studies to assess validity and reliability, and further develop these methods can be implemented by comparing updates or searches or expert selection for the same topics between different EPCs or other research groups.

With the electronic resources now available, ongoing updating could easily be carried out to avoid more lengthy and costly processes. Changes in situ can be made continuously, as the ACC hopes to do with its "updates" (versus full "revisions"). With constant perusal and the constant resources of a tabled group of experts, the guideline would never get far behind the curve. The original idea to set "full" review dates at specified intervals for each CPG would then regain credibility as the guideline would retain its authority during the interim.

The methods supporting and systematizing the development of clinical practice guidelines have evolved greatly over the last fifteen years. However, the
commitment to and study of guideline review, dissemination, and outcome evaluation has lagged. More research must be done in all three of these areas. Guideline review and update processes are gaining greater attention, but efforts need to be redoubled. Outdated guidelines are worse than no guidelines. Without further attention to this area of medicine, the promise of guidelines will not be realized, and, in fact, a potentially powerful means of improving patient care at the individual and societal levels may be undermined.

2 Eddy DM. Practice policies: where do they come from? JAMA. 1990; 263: 1269,72,75.

Other Works Cited

### Guideline/Update Information for Specific Specialty Organizations

<table>
<thead>
<tr>
<th>Organization</th>
<th>Contact Person</th>
<th>Has systematic update interval?</th>
<th>Has systematic process for deciding to update?</th>
<th>Has systematic process for doing update?</th>
<th>Has systematic process/criteria for expert selection?</th>
<th>Has farmed out portions of update process?</th>
<th>How long it takes to update?</th>
<th>Cost of updating?</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>Carla Herrerias, MPH</td>
<td>Yes, three years</td>
<td>No, has set schedule to update</td>
<td>Yes, full systematic review</td>
<td>No</td>
<td>Yes, works with AHRQ &amp; EPCs for syst. reviews</td>
<td>2 years</td>
<td>Rough approx. of $30,000 per update w/o incl EPC work</td>
</tr>
<tr>
<td>AACE</td>
<td>Sissy Crabtree, MPH</td>
<td>No, suggested &quot;expiration&quot; of 5 years</td>
<td>No, as dictated by intra-organization experts</td>
<td>No, can vary per workgroup chairman</td>
<td>No</td>
<td>No</td>
<td>12 months</td>
<td>Approx. $25,000 to write/publish Approx. $100,000 for total time involved</td>
</tr>
<tr>
<td>ACC</td>
<td>Paula Thompson, MPH &amp; Charlene May</td>
<td>No</td>
<td>Yes, focused lit. search and committee review each yr</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>12 months for update, 18-24 months for full revision</td>
<td>Unable to estimate</td>
</tr>
<tr>
<td>ACP-ASIM</td>
<td>Vincenza Snow, MD</td>
<td>No, but strict 5 year expiration date unless revision</td>
<td>Yes, ongoing focused search of lit. w/ committee review q 2 yrs</td>
<td>No, relies on contracted groups; method document being developed</td>
<td>Limited criteria includes CV, interview, &amp; publications</td>
<td>Yes, some evidence (sys) reviews by EPCs or other contracts</td>
<td>18 months</td>
<td>Unable to estimate</td>
</tr>
<tr>
<td>APA</td>
<td>Rebecca Thaler, MPH</td>
<td>No, suggested review of 5 years</td>
<td>No, as dictated by intra-organization experts</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Approximately 18 months</td>
<td>Very rough estimate of $50-75,000 per update</td>
</tr>
<tr>
<td>ASA</td>
<td>Richard Connis, PhD</td>
<td>No, approximate review time is 5 years</td>
<td>No, as dictated by intra-organization experts</td>
<td>Yes, full lit. review, meta-analysis, &amp; clin. consensus</td>
<td>No</td>
<td>No</td>
<td>2-3 years</td>
<td>Varying around $150-200,000</td>
</tr>
<tr>
<td>USPSTF</td>
<td>Russell Harris, MD, MPH</td>
<td>No</td>
<td>Yes, provisional method in place</td>
<td>No, although process under development</td>
<td>No</td>
<td>Yes, works with AHRQ &amp; EPCs</td>
<td>18-24 months</td>
<td>Rough estimate of $25-50,000</td>
</tr>
</tbody>
</table>
Figure 1.

Renewed ongoing systematic process for guideline review/update.
Back to methodologists

Methodologists/sifters provide new information every 3 months via literature review. (Experts queried annually)

Standing committee receives evidence quarterly. Meets annually to assess quality & importance of information.

New information forces 1 of 3 pathways

1) No changes required

2) Minor update/revisions required
Changes made by committee are incorporated after fulfilling required approval procedures.

3) Full update required

Note: every five years standing committee will formally review consolidated evidence to decide if full update is warranted.
Appendix One
Organizational Process For Developing A Clinical Practice Guideline

I. TOPIC IDENTIFICATION
- Section/Committee/Council/Task Force identifies a disease, condition, situation, or treatment (high risk; volume; diagnostic/treatment variability; availability of evidence; potential of legal liability; possibility of outcome modification; cost; target audience, etc)

II. DEVELOP INTENT FOR CLINICAL PRACTICE GUIDELINE

III. APPROVAL OF INTENT
- Steering Committee on Quality Improvement and Management
- Standing Board approval of Intent

IV. APPOINT ACADEMY SPECIALISTS, PRACTITIONERS, EPIDEMIOLOGISTS TO SUBCOMMITTEES
- Sections, Committees, Office-based Practitioners, Epidemiologists, outside medical organizations

V. BOARD OF DIRECTORS APPROVAL OF SUBCOMMITTEE MEMBERS VIA AAP EXECUTIVE COMMITTEE

VI. DEVELOP CLINICAL PRACTICE GUIDELINE
- In-depth literature review
- In-depth data analysis/meta-analysis
- Specific recommendations

VII. GATHER BASELINE DATA (optional)

VIII. REVIEW DRAFT CLINICAL PRACTICE GUIDELINE
- AAP Section/Committee Review
- Steering Committee on Quality Improvement and Management

IX. OUTSIDE ORGANIZATIONAL REVIEW

X. SEND TO MEDICAL EDITOR

XI. SEND TO BOARD OF DIRECTORS' REVIEWER

XII. APPROVED BY AAP EXECUTIVE COMMITTEE

XIII. DISSEMINATE AND IMPLEMENT CLINICAL PRACTICE GUIDELINES

XIV. REVISE THE CLINICAL PRACTICE GUIDELINE

V. DEVELOP CLINICAL PRACTICE GUIDELINE
- In-depth literature review
- In-depth data analysis/meta-analysis
- Specific recommendations
Clinical Practice Guideline Methodology

I. DEFINE THE PROBLEM
   • Select topic*

II. COMPARE INTERVENTION
    • Define; provider practice setting; classic case

III. IDENTIFY HEALTH OUTCOMES*
     • Patient preferences
     • Cost analysis
     • Health outcome measures

IV. DEVELOP STRUCTURE OF PROBLEM
    • Evidence model
    • Decision tree
    • Algorithm

V. DEFINE RESEARCH QUESTIONS
    • Critical friendships
    • Decision models

VI. CONDUCT LITERATURE REVIEW
    • Key Words
    • MeSH terms
    • Inclusion items

VII. DEVELOP/IMPLEMENT LITERATURE REVIEW FORMS

VIII. ABSTRACT THE EVIDENCE FROM THE ARTICLES

IX. DEVELOP EVIDENCE TABLES
    • Sensitivity analysis
    • Decision analysis

X. COMBINE THE EVIDENCE
    • Meta-analysis and clinical subjective judgement

XI. DEVELOP BENEFITS AND HAZARDS BALANCE SHEET

XII. SELECT THE BEST ALTERNATIVE STRATEGY**

XIII. DEVELOP CLINICAL PRACTICE GUIDELINE RECOMMENDATION

XIV. DISSEMINATE AND IMPLEMENT CLINICAL PRACTICE GUIDELINE

XV. REVISE THE CLINICAL PRACTICE GUIDELINE

*Think backwards. Start by identifying:
   1) health outcomes (length and quality of life; functional disability, death, etc)
      that patients experience and care about;
   2) interventions;
   3) definition of the problem, etc.

**Answer the question:
   Is this the best among the interventions, given the evidence,
   professional/clinical experience, cost, benefits, harms, etc, to achieve
   the specific health outcomes from the patient's perspective?
Appendix Two
AACE Standards on Medical Guidelines for Clinical Practice

1. **Purpose of the guideline is specified.**
   A. The guideline format should be clear and user friendly.
   B. The guideline should indicate how it is the same and how it differs from other guidelines covering the same area.
   C. The guideline should include a strategy for ongoing distribution.

2. **Rationale and importance of the guideline are explained.**
   A. A brief description of the review process that led to the development of the guideline should be given.
   B. The guideline should be specific enough so that it can be used in review articles as well as performance/quality measurements.

3. **The participants in the guideline development process and their areas of expertise are specified.**
   A. Describe the process for selection of the authors and reviewers of the guideline. List the credentials of these authors and reviewers.

4. **Targeted health problem or technology is clearly defined.**

5. **Target patient population is specified.**

6. **Intended audience or users of the guideline are specified.**

7. **The principal preventive, diagnostic, or therapeutic options available to clinician and patients are specified.**

8. **The health outcomes are specified.**
   A. The points of view put forth in the guideline must be supported by a scientific and clinical evidence and outcome data. The weight given to specific studies must be documented and the time frame of those studies must be indicated.
   B. The guideline must be comprehensive to the condition addressed and must include relevant outcomes of the interventions. For example, what is the likelihood of recurrence in a patient with a 1 cm capillary carcinoma of the thyroid if the patient has any hemithyroidectomy versus the total thyroidectomy.
   C. The guideline should indicate the outcome, which is the aim to be achieved in the application of the practice parameter (clinical condition). The outcome might include cure of the disease; delayed morbidity or mortality; symptomatic relief; improved quality of life; conservation of resources (prevention of antibiotic resistance, for example), impact on patient and physician behavior.
D. The guideline should indicate the measures used to determine those outcomes. For example, length of life, quality of life, length in the hospital, efficiency of control of overhead factors.

9. **The method by which the guideline underwent external review is specified.**

10. **An expiration date or date of scheduled review is specified.**
   A. The guideline should be reviewed and revised probably within five years or if new research findings emerge.
   B. The guideline should specify the planned review date.
   C. The guideline should indicate that a system is established to monitor the emergence of the information, which may necessitate revision.
   D. The guideline should include the methods to evaluate its own outcome performance and specifically to measure the guideline’s impact and affect and indicate how this information will be used to revise the updated guideline.

11. **Method of identifying scientific evidence is specified.**
   A. Explain the rationale for including or excluding studies related to the topic.
   B. Indicate in the methods used to evaluate the scientific literature and other research findings. Explain why articles are included or excluded. List all the articles included in the guideline.

12. **Time period from which evidence is reviewed is specified.**
   A. The guideline literature review should be current within 3 to 5 years.

13. **The evidence used is identified by citation and referenced.**
   A. The guideline should be clearly referenced in a standard scientific paper format.

14. **Method of data extractions is specified.**

15. **Method for grading or classifying the scientific evidence is specified.**
   A. The guideline must indicate if the expert opinion given by the authors of the guidelines was linked to scientific evidence (to create their recommendations).
   B. The guideline can and should incorporate findings from review of the literature and from the clinical judgment of the authors of the guideline. The weight given to each type of finding should be described.
C. Outcome data, when available, should be published also and the process by which the particular outcome data were selected should be described.

16. Formal methods of combining evidence or expert opinion are used and described.
   A. If the recommendation of the guideline authors differs from the data in the published literature, an explanation for this difference must be given.

17. Benefits and harms of specific health practices are specified.
   A. The guideline should indicate the appropriateness of the recommendations in specific clinical circumstances. The absolute nature or flexibility of the recommendation should be indicated. For example, if there are different ways to treat a particular condition, the rationale for the best treatment should be given and the rationale for the best alternative treatments should also be given.
   B. The guideline should provide the information (clinical, scientific, and economic) needed to make decisions.

18. Benefits and harms are quantified.

19. The effect on health care costs from specific health practices is specified.
   A. The guideline should be adaptable to different practice environments such as private offices, hospitals, clinics, and other health care facilities.
   B. Where appropriate economic data should be given. The method for collecting and evaluating these economic data should be described. For example, in the treatment of hyperthyroidism indicate the cost of radioiodine, surgery, and medication in different parts of the country and describe how this information was obtained.

20. Costs are quantified.
   A. The guideline should note the cost of one procedure over another and indicate how that affects patient care and outcomes.

21. The role of value judgements used by the guideline developers in making recommendations is discussed.

22. The role of patient preference is discussed.
   A. Where appropriate the guideline should define and assess patient preference data. For example, two patients prefer one form of treatment of hyperthyroidism over another and if so explain why and whether this patient choice is compatible with good care and cost effectiveness.
B. The guideline should provide data regarding patient preference for one treatment over another and explain in the rationale for that preference (cultural, regional, alternative medicine approach).

23. Recommendations are specific and apply to the stated goal of the guideline.
   A. Recommendations are graded according to the strength of the evidence.
   B. The guideline should clearly state limitations of the recommendations in the guideline and should state the process used to formulate those limitations.

25. Flexibility in the recommendations is specified.

References

Appendix Three
Manual for ACC/AHA Guideline Writing Committees

Methodologies and Policies from the ACC/AHA Task Force on Practice Guidelines

Section I: Overview of Methodology and Purpose of the Manual

Importance of ACC/AHA Guidelines
The creation of clinical practice guidelines has been a joint activity between the American College of Cardiology (ACC) and the American Heart Association (AHA) since the 1980s. The guidelines advance the missions of both organizations by providing clinical recommendations to health care providers for the purpose of improving cardiovascular health. The Institute of Medicine defines practice guidelines as, “systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances.” (1990) Well developed guidelines have the potential to enhance the appropriateness of clinical practice, improve the quality of cardiovascular care, lead to better patient outcomes, improve cost-effectiveness, and identify areas of further research needs.

Purpose and Scope of the Manual
To continue as a leader in the field of clinical practice guidelines, the ACC/AHA Task Force on Practice Guidelines (Task Force) has overseen the creation of this manual to assist guideline writing committees in navigating guideline creation. The bulk of this manual consists of tools to assist guideline writers in interpreting and applying the methodology. A flowchart highlighting the key steps in the development of evidence-based guidelines (Figure 1) serves as the basis for organizing the manual.

The Task Force understands the challenges in applying a uniform methodology to guidelines that represent diverse diseases, conditions, diagnostics, and interventions. In all cases, writing group members should familiarize themselves thoroughly with the
manual, as these policies and standards provide the framework for guideline creation. However, if warranted the Task Force may allow exceptions to the written policies.

**Staff Support**
The ACC and AHA provide scientific and administrative staff to support the creation of evidence-based guidelines. A Research Analyst and Guideline Manager are assigned to each guideline to assist writers with the methodology and process of guideline development.

© Back to top

Page last updated 03/12/2002
Figure 1. Steps in the Development of Evidence-Based Guidelines

**Step One**
Determine the guideline scope and clinical objectives

**Step Two**
Define and conduct appropriate and comprehensive literature searches

**Step Three**
Sort and evaluate the evidence

**Step Four**
Synthesize and interpret the evidence

**Step Five**
Write recommendations based on expert interpretation of the evidence

**Step Six**
Assign classification of recommendations and strength of evidence

**Step Seven**
Create tables, diagrams, and mnemonics describing recommendations
Manual for ACC/AHA Guideline Writing Committees

Methodologies and Policies from the ACC/AHA Task Force on Practice Guidelines

Section II: Tools and Methods for Creating Guidelines

- **Step One:** Determine the Guideline Scope and Clinical Objectives

**Topic Selection**

ACC/AHA clinical practice guidelines are written on three general categories: health conditions, procedures, and diagnostics. The Task Force determines the topics for guidelines and selects the writing committee members, while the writing committee is responsible for developing the guideline's content.

**Determining the Guideline's Scope**

Before and during the first meeting, the writing committee primarily focuses on coming to consensus about the guideline's scope (see Checklist 1). Literature searching is conducted to determine the scope of the guideline, as appropriate (see Step Two).
ACC/AHA guidelines are usually intended to provide recommendations applicable in the United States; however, some guidelines written in collaboration with the European Society of Cardiology or other partners have a broader target audience. The methodology for international guidelines is the same as national guidelines, with conclusions and recommendations based on expert judgement applied to clinical evidence. International differences in disease management and health care resource availability may be noted when such differences might have significant impact on the implementation of recommendations.

Although some guidelines also address issues of cost-effectiveness and related economic analyses, ACC/AHA guidelines are generally meant to provide clinically relevant information outside of the context of costs and reimbursement. If cost issues must be included, guideline writers should limit the scope to previously published analyses and not attempt to create any new economic analysis within the document.

Guideline Updates and Revisions
Maintaining guideline content that is up-to-date with the clinical evidence and best practices in the field of cardiology is an ongoing challenge. The Task Force is exploring new systems to update guideline content more regularly. In the meantime, all guidelines are reviewed for possible update within one year after publication. The research analyst and the chair monitor literature on the topic, and compare the current guideline recommendations against the latest data. The entire writing committee is periodically surveyed to determine if the guideline (or sections within the guideline) needs updating. Guideline updates should focus on substantive changes to recommendations rather than editorial changes to the document. Otherwise, all methodology in this manual applies to updates and revisions.

Guideline Structure
Guideline writers are encouraged to define as precisely as possible the overall guideline structure at the early stages of guideline creation. The Task Force has provided standard guideline outlines for each guideline type (see Table 2). These outlines improve consistency across guidelines and facilitate the effectiveness of on-line searching of our guidelines. They provide a common structure while allowing for flexibility as the topic demands. Guideline writers should determine the outsets which "standard concepts" apply to their guideline, then proceed with creating detailed clinical objectives under each concept. The standard outlines are not prescriptive, nor are they meant to encourage the creation of textbook-style guidelines.

Identifying the Clinical Objectives
The main goal of guideline creation is to develop recommendations that allow providers to understand the evidence on the topic and apply it to clinical practice. As such, guideline writers should progress with specific clinical objectives in mind. It may be very helpful at the outset to consider what kind of guidance the readers will expect in the completed document, such as:

- The role of exercise testing in asymptomatic patients

Ideally, a set of guidelines should give practitioners, patients, and policymakers an explicit description of the projected health benefits and the projected harms or risks. In addition, projected outcomes should be compared with those for alternative courses of care for the clinical condition in question.

-Institute of Medicine, 1990
The use of inotropic agents in patients with end-stage heart failure
- Managing mitral regurgitation medically versus surgically

A comprehensive collection of clinical objectives should be created within each main concept addressed by the guideline outline. These clinical objectives serve as the basis for literature searching and sorting, and later for the compilation of guideline recommendations.

Checklist 1. Determining the Guideline Scope and Clinical Objectives

Questions related to the guideline overall

- What is the guideline's targeted health condition, procedure, or diagnostic?
- What is the purpose of the guideline?
- What is within the scope of the guideline?
- What is outside the scope of the guideline?
- What is the epidemiology of the topic?
- Who are the guideline's intended users?
- What is the target patient population to be addressed in the guideline?
- How does the guideline relate to other existing ACC/AHA guidelines?
- Can a few flow diagrams summarize the guideline, or at least key sub-sections?

Questions related to the guideline's clinical objectives

- What are the important clinical objectives related to the guideline topic?
- What sub-topics and related topics must be included in the guideline?
- Are flow diagrams appropriate to these sub-topics and related topics?
- What are the potential benefits and risks for individual patients associated with an intervention or procedure?
- What amount of clinical flexibility is appropriate for the topic area?
- What clinical options are available?
- What topics have already been covered in existing ACC/AHA guidelines?
Table 2. Standard Guideline Outlines

### Disease or Condition Guidelines

<table>
<thead>
<tr>
<th>Standard Concepts</th>
<th>Possible Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>Purpose of the guideline Scope</td>
</tr>
</tbody>
</table>
| Definition of the disease/condition | Overview  
Epidemiology  
Classifications  
Characterizations               |
| Clinical Evaluation          | Recognition  
Methods for Risk Stratification  
Other Issues Related to Clinical Assessment |
| Diagnosis and Testing        | Non-invasive testing  
Invasive testing  
Laboratory testing  
Risk Assessment               |
| Treatment                    | Principles of Management  
Therapy  
Medication  
Procedures  
Interventions  
Alternative/complimentary medicine  
Monitoring                     |
| Special populations          | Concomitant disorders  
Patient groups                   |
| Follow-up                    | Discharge  
Long-term management            |

### Intervention Guidelines

<table>
<thead>
<tr>
<th>Standard Concepts</th>
<th>Related content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>Purpose of the guideline Scope</td>
</tr>
</tbody>
</table>
| Definition of intervention/procedure | General considerations  
Background                       |
| Indications                  | Clinical uses                                                                    |
| Management strategies        | Procedure-specific considerations  
Associated medical therapies  
Procedural complications  
Reducing risk                   |
| Outcomes                     | Definitions of success  
Short-term and long-term outcomes  
Comparisons with other interventions |
| Institutional/Operator issues | Quality assurance/improvement  
Volume considerations            |
| Special populations          | Patient groups                                                                   |

### Diagnostic Guidelines

<table>
<thead>
<tr>
<th>Standard Concepts</th>
<th>Related content</th>
</tr>
</thead>
</table>
| Introduction                 | Purpose  
Scope                     |
| Description of the diagnostic tool | Specific procedures  
Equipment  
Sensitivity/specificity  
General considerations  
Comparison with other diagnostic tools |
| Specific conditions          | Clinical uses (note: diagnostic guideline usually subdivided by the disease/conditions that they can diagnose. These discussions included diagnosis, assessment, prognosis, risk stratification, screening, etc.) |
| Special populations          | Patient groups                                                                   |
Manual for ACC/AHA Guideline Writing Committees

Methodologies and Policies from the ACC/AHA Task Force on Practice Guidelines

Section II: Tools and Methods for Creating Guidelines

- **Step Two: Define and Conduct Appropriate and Comprehensive Literature Searches**

Finding and Managing the Evidence

Once the scope of the guideline has been determined, comprehensive searching of the published literature takes place. A key component of the ACC/AHA guideline methodology is the creation of recommendations based on the entirety of the evidence currently available. The Institute of Medicine describes literature searching as the key step in developing valid guidelines.

It has been estimated that over 2 million articles and more than 17,000 biomedical books are published annually. The challenge of finding relevant articles among the millions is compounded by the availability of multiple electronic databases, all of which offer different but partially overlapping pools of information.
The current resources for ACC/AHA guideline development allow for searching in MEDLINE (via PubMed), EMBASE, the Cochrane Library, and Best Evidence. The Research Analyst assigned to the guideline will manage a computerized database of all citations relevant to the guideline topic. The Research Analyst and Librarian conduct searches and forward relevant citations to the writers.

**Literature Search Methodology**

Figure 2 outlines the ACC/AHA process for conducting comprehensive literature searches for the guidelines. Initial literature searching focuses on published meta-analyses and systematic reviews. If high quality, relevant, and up-to-date meta-analyses or systematic reviews are found, these articles allow writers to focus on critiquing and updating an existing review as opposed to creating one. For the majority of topics, literature searching also includes randomized clinical trials, and is expanded to non-randomized studies, case studies, and opinion documents until the evidence-base is sufficient for each clinical question identified in Step One. Each article should be critically evaluated as to quality and clinical limitations, as discussed in Step Three.

**Documentation of Searching**

All literature searching for guideline development must be documented by the searcher and stored at the ACC offices. This allows the chair and Research Analyst to construct the text of the guideline describing the literature search criteria, thereby allowing guideline users to assess the comprehensiveness of the searching.

In addition to searches conducted by staff, writing committee members are welcome to conduct their own literature searches, including search criteria beyond what the ACC/AHA resources are able to provide (see below: Standard Search Criteria for ACC/AHA Guidelines). The documentation for all literature searches must be forwarded to the Research Analyst using the Template: Literature Searches for ACC/AHA Guidelines included in this section.

**Standard Search Criteria for ACC/AHA Guidelines**

- Literature searching includes the following on-line databases:
  - MEDLINE/PubMed
  - EMBASE
  - Cochrane Library
  - Best Evidence
- Searches are limited to English language. (Searches will be expanded to languages other than English on a case-by-case, as requested basis.)

Because a guideline based on an incomplete or biased evaluation of the literature can lead to inappropriate recommendations, the search for relevant research should be comprehensive, research should be selected using explicit criteria, and the validity of the results should be judged in a rigorous and reproducible fashion.

- Cook, 1997
• Searches are limited to human subjects.
• In the case of a guideline update, searches are limited to the time period following the publication of the last version of the guideline.
• In the case of a new guideline or full revision, no time limits on searches are imposed, unless the writing committee determines that a different time frame appropriate (for example, a guideline on a diagnostic that did not exist before a certain date).
• Gender and age are not limited, except when a specific clinical objective applies only to a particular sex or age group.
• Publication type is initially limited to meta-analyses and systematic reviews. Publication type is expanded on an as-needed basis to include randomized controlled trials, non-randomized studies, case studies, and opinion documents.
• If an acceptable systematic review or meta-analysis is identified, searches to update it are typically limited to the time period following the search cut-off date reported in the review.
Figure 2. Process for Conducting Comprehensive Literature Searches for Guidelines

Guideline writer forwards search requests to Research Analyst

Librarian searches for meta-analyses and systematic reviews

Research Analyst removes non-relevant articles and sorts relevant articles by clinical question or objective

Are there high quality, relevant, and up-to-date meta-analyses or systematic reviews that answer the identified questions?

Yes

Librarian searches for randomized controlled trials

Research Analyst removes non-relevant articles and sorts relevant articles by sub-topic/guideline section

Are there high quality, relevant, and up-to-date randomized controlled trials that answer the identified questions?

Yes

Librarian searches for non-randomized studies

Research Analyst removes non-relevant articles and sorts relevant articles by sub-topic/guideline section

Are there high quality, relevant, and up-to-date non-randomized studies that answer the identified questions?

Yes

Librarian searches for case studies and opinion documents

Research Analyst removes non-relevant articles and sorts relevant articles by sub-topic/guideline section

No

No

No

Proceed to Step Three
Manual for ACC/AHA Guideline Writing Committees

Methodologies and Policies from the ACC/AHA Task Force on Practice Guidelines

Section II: Tools and Methods for Creating Guidelines

- **Step Three: Sort and Evaluate the Evidence**

**Stages of Sorting Evidence**

After the literature search results have been imported into the computerized database managed at the ACC, the Research Analyst reviews the abstracts and removes non-relevant citations. At this step, only the article's title and abstract are assessed, so any article likely to be relevant to the guideline is maintained. Additionally, the Research Analyst sorts the abstracts to correspond with the specific clinical objectives identified in Step One. This initial sort creates a comprehensive set of potentially relevant studies.

Although the Research Analyst does a preliminary level of sorting, the clinical expertise of writing committee members is necessary to make the final decision as to whether the article is a relevant piece of evidence that should be included in the development of a recommendation. This often requires review of the article's full text and critique of the research methodology employed. As necessary, the research analyst will provide the full text of all peer-reviewed, published:

- Randomized controlled trials
- Meta-analysis
- Systematic reviews of evidence
- Diagnostic studies using comparison with a gold standard

*It is important to err on the side of over-inclusion because once a trial has been excluded from the selection process it is unlikely to be reconsidered. Questionable articles which are included at one stage can be excluded at a latter stage when more information on the study is available.*

- Mulrow, 1996
Along with each full text article, the writer will receive Checklist 2, which asks the writer to make the final determination of the article's role within the guideline. Documentation of completed checklists will be maintained by the ACC in the methodology files for the guideline. The Appendix includes three articles from the British Medical Journal and articles by Pogue and Lau that provide further information on evaluating the quality of published literature.

Unpublished Data
Guideline writers are frequently familiar with data from abstracts and late breaking trials that may impact the guideline's content. The results from unpublished data should not be considered except in few instances, should be no older than 2 years, and should be clearly stated to be unpublished data in the guideline text. Only trials presented at a major national or international scientific meeting are allowed, and may not be used to support any recommendation. When trial data are discussed, the text should clearly state that the data are preliminary. Additionally, guideline writers should obtain slides from the trial presentation, perform a detailed review, and ask the presenter of the trial for guidance, keeping in mind that the trial group has the prerogative to request that the information not be published in a guideline.

Balancing Scientific Rigor with Feasibility
The Cochrane Collaboration publishes perhaps the most rigorous and comprehensive guide to conducting systematic reviews of evidence, and their methodology has provided the basis for much of this manual. However, due to time and economic constraints, some components of their methodology (such as creating and validating criteria for which articles to include, and removing the journal and author names from articles being reviewed) are beyond the scope of ACC/AHA guideline development.

A less resource-intensive, more feasible approach is to establish a few basic criteria (such as randomized controlled trials only or studies with at least six month follow-up) and be as inclusive and unbiased as possible. The Task Force recommends rigorous review of the articles used in evidence tables and meta-analyses—those articles that are most fundamental to the guideline recommendations. Documentation of why studies are included and excluded from consideration will provide additional scientific rigor to the document and will be published on the ACC web site as a component of the guideline methodology. The Research Analyst assigned to the guideline can and should be used in scoring the articles and synthesizing the evidence (Step Four) to ease the burden on the writing committee.
Checklist 2. Determining the Evidence-Based for Guideline Recommendations*

Guideline writers are asked to consider the merits, quality, and generalizability of each article relevant to the clinical objective. This checklist should be completed only for articles from the peer-reviewed, published literature that are:

- randomized controlled trials,
- meta-analysis/systematic reviews, or
- diagnostic studies using comparison with a gold standard.

Article/Author:
Please indicate one of the following conclusions about the article:

_____ Yes, this is a relatively high quality study that provides credible results and should be included in the evidence table and references that support the recommendation(s) for this clinical objective.

_____ No, this study is not of sufficient quality to be included in the evidence table, but Yes, this study contains some useful information about the clinical objective and should be maintained as a reference for the text accompanying the recommendation.

_____ No, this is a relatively poor study that should not be used in the evidence table or in the references for this clinical objective.

_____ No, this study is not directly relevant to the clinical objective.

Comments:

* Our methodology for grading individual studies is a work in progress. We are currently pilot testing this checklist with a number of writing groups as well as investigating more detailed checklists for evaluating study quality. Please re-visit this site in the coming months for newer checklists.
Section II: Tools and Methods for Creating Guidelines

- **Step Four: Synthesize and Interpret the Evidence**

**Guideline Authoring Template**
To improve the consistency of guideline content, both within and between guidelines, the Task Force has created a guideline authoring template. Guideline writers will receive one template for each clinical objective they are responsible for writing. The following is a brief introduction to using the template:

- The fields for **guideline**, **author**, **section name**, and **clinical objective** will be completed by staff.
- The **evidence base** field will consist of the articles gathered through the literature search and sorting (Steps Two and Three). Writers whose clinical objectives have a large evidence base may also receive full text articles or an evidence table along with the template.
- There are four fields for **recommendation**, including a checklist for classification and level of evidence for each recommendation (more than four recommendations can be written). After reviewing the evidence base, the writer should create recommendations that answer the clinical objective. See **Steps Five and Six** for details on writing recommendations.
- The **text** field is used for placing recommendations in context (see "Narrative synthesis of evidence" below).
- The **additional references** field can be completed by a writer who references other sources than those provided in the evidence base field.
- If the clinical objective has a **diagram**, **table**, or **graphic** associated with it, it can be added or referenced in this field (see **Step Seven**).
Clinical Statements / Manual - Section II

Narrative Synthesis of Evidence
Summaries of evidence should generally be in tabular form, and not in the text of the guideline. Text should be reserved for qualifying or clarifying the recommendations. The Task Force prefers that clinical trial data and other evidence be displayed in an evidence table or included in meta-analysis. When multiple trials have yielded similar, non-controversial results (e.g., the use of aspirin post myocardial infarction), a single sentence with appropriate references may suffice. Long, descriptive paragraphs of the methodology and findings of individual trials are discouraged.

Visual Synthesis of Evidence
Preparing an evidence table involves identifying and extracting the key data from the relevant studies. The Cochrane Collaboration recommends beginning by deciding what comparisons need to be made, then identifying the data elements necessary to make those comparisons. Salient data elements may include, but are not limited to, number of patients, morbidity, mortality, dose-response, sensitivity, specificity, p-values, confidence intervals, positive predictive value, negative predictive value, and relative risk.

The next step is to prepare visual summaries of the results of the studies included in each comparison. The data are usually displayed in a table that allows the studies' designs and results to be easily compared. However, sometimes the data are better summarized in a bar chart or other graphic summary. Information presented graphically can replace the need for "text-heavy" sections of the guideline. Examples of visual synthesis of evidence from published ACC/AHA guidelines include:

- Acute Myocardial Infarction Guideline, Figure 7
- Atrial Fibrillation Guideline, Figure 16
- Chronic Stable Angina Guideline, Table 21

Analytical Synthesis of Evidence
Sometimes recommendations can confidently be written based on the organization of evidence in tables or graphs. Other times, a further step is necessary; analyzing the data statistically to get an estimate of the heterogeneity of the individual effect sizes, an estimate of the summary effect size, and a measure of its variance. Guideline writers generally rely upon meta-analytic methods to conduct such analyses.

A detailed guide to the methods of meta-analysis is beyond the scope of this manual. However, ACC staff is available to assist writing committees in conducting meta-analysis for guidelines.

Also recommended is the Cochrane Collaboration, which offers free on its web site (www.cochrane.org) software called Review Manager 4.1. RevMan allows for entry of the characteristics of studies and their findings, and the creation of comparator tables. It can perform meta-analysis of the data entered, and present the results graphically. A comprehensive handbook for conducting systematic reviews that can be printed from the same web site accompanies the software.
Use of Other Guidelines/Authorities
Guideline text, recommendations, and evidence tables may be replicated from previous ACC/AHA guidelines and statements endorsed by both organizations (e.g. National Cholesterol Education Program). Consensus statements or guidelines developed by others and not endorsed by the AHA and ACC should not be cited or referenced unless absolutely necessary, as this implies endorsement on the part of the organizations.

Discussing Pharmacotherapy in Guidelines
The Task Force has provided a detailed list of policies on discussing pharmacotherapy in guidelines as Checklist 3. In addition, a pharmacologist will either be assigned to each guideline or will be used in a consulting role to review the guideline’s pharmacotherapy discussions before publication.

Investigational treatments or drugs that are not available for general use may be mentioned, but should be clearly described as such and not given Class I, IIa, or III recommendations. The writing committee should decide whether to list them as Class III, or not to list them at all. The presence or absence of FDA or CMS approval of a drug or device for a specific purpose should generally not be mentioned. The criteria used by regulatory authorities are frequently different, and the ACC/AHA process should be independent of these regulatory issues.

In the case of international guidelines co-sponsored by the ACC/AHA/ESC, it is understandable that rare occasions may require a discussion of international availability of certain medications. However, such content should be addressed from the perspective of the patient or indication, and not from a policy (i.e. drug approval) perspective.
Checklist 3. Discussing Pharmacotherapy in Guidelines

- Use generic or chemical name not trade name
  - e.g., simvastatin, not Zocor
- Use broadest and most generic name of class appropriate
  - e.g., cholesterol-lowering not "statins"
- List classes of drugs or drugs within classes according to evidence-based rationale and state rationale
  - e.g., first-line, second-line or side effects or cost-effectiveness
  - If no evidence-based rationale, list alphabetically
- List all drugs (or none) within class
  - Indicate whether each is approved for the indication(s) under discussion
    - e.g., statins for primary prevention
  - Indicate whether each has evidence for the indication(s) under discussion
    - e.g., llb/llla’s
- Discuss evidence for or against "class effect"
  - e.g., issue raised by ramipril in HOPE study
- When so-called "alternative medicines" are known to be widely used, discuss the evidence about them and the issues raised by their use
  - e.g., possible interactions
- Avoid the use of symbols and abbreviations when discussing drug dosing and timing.
  - e.g., use "micrograms" or "mcg" instead of "Fg"
  - The Institute for Safe Medication Practices has issued a drug error alert regarding some commonly used abbreviations
- Whenever a guideline includes specific drug information, such sections of the guideline should be reviewed by a pharmacologist during peer review.

© Back to top

Page last updated 03/12/2002
Step Five: Synthesize and Interpret the Evidence

Recommendations: The Essence of Guidelines
Steps One through Four describe the methodology of guideline development as well as meta-analysis and systematic review. Guideline development, unlike the other methodologies, goes beyond the compilation and analysis of data to include recommendations that guide clinical practice. Guideline writers are challenged with considering a vast array of evidence and creating clinically applicable and clear recommendations from it.

As the evidence is considered, conclusions and recommendations naturally evolve. Whenever this occurs, the recommendation should be condensed into a sentence or two and separated from the text (see Step Six). The recommendations are the core guideline content, while the text enhances the recommendations by providing further descriptive information, such as exceptions to the recommendations and clinical options.

If Step One determined that flow diagrams were appropriate, recommendations should be incorporated into the flow diagrams where appropriate (see Step Seven).

Because guidelines are increasingly serving as the basis for other ACC and AHA activities (such as pocket guides, performance measures, data...)

It does not matter how much information you have or how valid and relevant it is, It should always be modulated by the values and preferences of the decision makers and those who will be affected by the decisions ...

In most cases, however, it will be unclear how much your anecdotal information, values and preferences should modulate the research evidence available.

- Jadad, 1998
standards, and GAP Projects), recommendations should be stand alone text that are written in complete sentences with as much detail as possible. Guidelines are intended to be applied by health care providers in real world settings, so the recommendations should be practical, feasible, and clinically flexible, thus facilitating the translation and implementation of recommendations.

Expert Interpretation of the Evidence
ACC/AHA clinical practice guidelines are written by cardiologists, other experts in the field of cardiology or cardiovascular research, and representatives from other organizations and specialties, when relevant. Their scientific and clinical expertise is germane to the creation of guideline recommendations that are useful to a broad spectrum of health care practitioners.

Despite all the evidence that may be available for writing the guideline, expert interpretation will always be necessary. Expert interpretation serves as a funnel through which evidence on multiple questions and clinical situations is combined, condensed, and formulated into recommendations.

Unfortunately, most evidence falls into the "gray zone" of uncertainty. The evidence from different trials may come to divergent conclusions, the evidence may only apply to specific sub-populations, the evidence may be from methodologically weak studies, or the evidence may simply be insufficient to make a decision. Only in rare instances is there an abundance of evidence available that leads directly to an indisputable recommendation.

Writing guidelines and formulating recommendations are not simple tasks. The guideline writer is frequently in a dilemma as to whether to delay making a decision or come to a conclusion despite the holes in the evidence. Checklist 4 (see below) in this section provides a list of qualities of guideline recommendations to consider when writing the document.
Checklist 4. Writing Guideline Recommendations

- Write all recommendations in complete sentences.
- Write separate recommendations that apply to specific clinical objectives.
- Write recommendations that are practical in the real world setting.
- Describe the patients to whom the recommendation applies.
- Use unambiguous language and clearly defined terms when writing recommendations.
- Write recommendations in terms of active/positive actions rather than passive/negative actions, e.g., Class I recommendation to perform a test/give a treatment that is useful/effective rather than a Class III recommendation not to perform/give it.
- When there are areas of uncertainty or controversy include this information in the recommendation.
- Quantify as much as possible benefits, harms, and timeframes.
- Write recommendations that incorporate data on patient preferences, when applicable.
- Specify sub-population variability and exceptions in the recommendations. List the exceptions whenever possible.
- Include flexibility in applying the recommendations, where applicable.
- Recommendations must be consistent with previous ACC/AHA guidelines, unless there is new evidence to justify a change. Both the new evidence and the change must be described in detail.
Once recommendations are written, a Classification of Recommendation and Level of Evidence grade must be assigned to each recommendation. Classification of Recommendations and Level of Evidence are as follows:

**Classification of Recommendations**

**Class I:** Conditions for which there is evidence and/or general agreement that a procedure or treatment is useful and effective.

**Class II:** Conditions for which there is conflicting evidence and/or a divergent opinion about the usefulness/effectiveness of a procedure or treatment.

**IIa:** Weight of evidence/opinion is in favor of usefulness/effectiveness.

**IIb:** Usefulness/effectiveness is less well established by evidence/opinion.

**Class III:** Conditions for which there is evidence and/or general agreement that a procedure/treatment is not useful/effective, and in some cases may be harmful.

**Level of Evidence**
Clinical Statements / Manual - Section II

Level of Evidence A  Data derived from multiple randomized clinical trials
Level of Evidence B  Data derived from a single randomized trial, or non-ran
Level of Evidence C  Consensus opinion of experts

Applying the Classifications and Levels
Some writers prefer to assign the Classification of Recommendation and Level of Evidence when writing the recommendations, while others prefer to state the recommendation and assign the classification later after re-examining the data. Writers preferring the first method will conduct Steps Five and Six of the guideline methodology simultaneously.

The Classification of Recommendations and Level of Evidence are considered by many to be the core of the guidelines. As such, they are among the most debated aspects of the guideline within the writing group. See the section on writing committee consensus development for guidance on coming to agreement on recommendations.

Any combination of Classification of Recommendation and Level of Evidence is possible. For example, a recommendation can have a Class I, even if it is entirely on expert opinion and no research studies have ever been conducted on the recommendation (Level C). Similarly, a Class IIA or IIB can be assigned a Level A if multiple randomized controlled trials coming to divergent conclusions.

Assigning a Level of Evidence B or C should not be construed as implying that the recommendation is weak. Many important clinical questions addressed in the guideline do not lend themselves to experimentation or have not yet been addressed by high-quality investigations. Even though randomized controlled trials may not be available, the question may be so relevant that it would be delinquent to not include it in the guideline.

Table 3. Applying Classification of Recommendations and Level of Evidence
<table>
<thead>
<tr>
<th>Level</th>
<th>Class I</th>
<th>Class IIa</th>
<th>Class IIb</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>o Recommendation that procedure or treatment is useful/effective</td>
<td>o Recommendation in favor of treatment or procedure being useful/effective</td>
<td>o Recommendation's usefulness/efficacy less well established</td>
<td>o Recommen</td>
</tr>
<tr>
<td>A</td>
<td>o Sufficient evidence from multiple randomized trials or meta-analyses</td>
<td>o Conflicting evidence from multiple randomized trials or meta-analyses</td>
<td>o Conflicting evidence from multiple randomized trials or meta-analyses</td>
<td>d for not use</td>
</tr>
<tr>
<td>B</td>
<td>o Recommendation that procedure or treatment is useful/effective</td>
<td>o Recommendation in favor of treatment or procedure being useful/effective</td>
<td>o Recommendation's usefulness/efficacy less well established</td>
<td>o Recommen</td>
</tr>
<tr>
<td></td>
<td>o Insufficient evidence from single randomized trial or non-randomized studies</td>
<td>o Conflicting evidence from single randomized trial or non-randomized studies</td>
<td>o Conflicting evidence from single randomized trial or non-randomized studies</td>
<td>d for not use</td>
</tr>
<tr>
<td>C</td>
<td>o Recommendation that procedure or treatment is useful/effective</td>
<td>o Recommendation in favor of treatment or procedure being useful/effective</td>
<td>o Recommendation's usefulness/efficacy less well established</td>
<td>o Recommen</td>
</tr>
<tr>
<td></td>
<td>o Only expert opinion, case studies, or standard-of-care</td>
<td>o Only diverging expert opinion, case studies, or standard-of-care</td>
<td>o Only diverging expert opinion, case studies, or standard-of-care</td>
<td>d for not use</td>
</tr>
</tbody>
</table>

Table
Section II: Tools and Methods for Creating Guidelines

- Step Seven: Create Tables, Diagrams, and Mnemonics Describing Recommendations

Once the evidence tables and recommendations have been created, guideline writers should look for ways to visually summarize the key points in tables, diagrams, and mnemonics. The flow diagrams identified in Step One should be considered again in light of the evidence collected and recommendations written. Frequently, the text and/or recommendations can be condensed into a clinical pathway, algorithm, or decision-tool. These visual summaries assist physicians in understanding and applying the best care for individual patients. Visual presentations should be:

- Written in clear and unambiguous language.
- Logically organized.
- Easy to follow.
- Specific about relevant populations and clinical circumstances.
- Specific about which elements of care are appropriate, inappropriate, and equivocal.

The broad mandate of most guidelines ensures that guideline documents tend to be longer and less formulaic than other articles... [but] surveys reveal that clinicians prefer pocket cards, concise pamphlets, and journal article summaries.

- Cook, 1999

Note: These documents are best viewed with Internet Explorer version 5.0 and higher or Netscape Navigator 6.0 and higher.
The guideline users will expect the evidence to be presented as proof of the recommendations' quality. However, in clinical circumstances, the key points of how the evidence applies to patients are the take-home messages that must be clearly presented and easily accessible in the guideline.

Examples of good tables, diagrams, and mnemonics, include:

- Atrial Fibrillation Guideline, Figure 9
- Chronic Stable Angina Guideline, Figure 5
- Perioperative Evaluation Guideline, Figure 1

Preparing the Guideline for the Pocket Guide

Many of the ACC/AHA guidelines are converted into a pocket guide version to facilitate implementation of the guideline, specifically at the point of care. The information in the pocket guide should flow directly from the full guideline; thus guideline writers are responsible for ensuring that the guideline lends itself to the pocket guide format. The Task Force recommends that each writing committee designate one writer who will be responsible for the pocket guide.
Writing Committee Discussions and Consensus Development

Referring back to Figure 1, writing committee discussions and consensus development is ongoing at all stages of guideline development. Since ACC/AHA guidelines are team-written documents, coming to agreement on the scope, clinical objectives, evidence tables, text, recommendations, and visual summaries occurs throughout document development. Sub-section writers often come to consensus through phone calls or e-mail exchanges of information, while the entire writing committee comes to consensus during the face-to-face meetings.

In evidence-based documents such as clinical practice guidelines, consensus development is often most important around topics that have no literature base. Writing groups are faced with the challenge of addressing an important clinical question despite a lack of data. The ACC/AHA guideline development process allows for the incorporation of minority opinions within the document if consensus can not be reached.

Finalizing the Document
At the final stages of guideline development, writers should re-examine the original goals regarding the scope of the guideline, as identified in Step One. Any identified gaps should be filled or explained before the document is sent to peer review. The writing group will be asked to give formal approval of the document.
both before peer review and after peer review edits have been incorporated. Checklist 5 (see below) is provided as a tool to conduct an internal review of the guideline recommendations at both of these junctions.

Checklist 5. Reviewing Guideline Recommendations

- Are the recommendations within the stated purpose and scope of the guideline?
- Are all recommendations cited and referenced (either in the text or in the evidence table)?
- Are all recommendations assigned a Classification of Recommendation and a Level of Evidence?
- Are clinically important and feasible recommendations made?
- Are areas of uncertainty and exceptions to the rule clearly identified?
- Are evidence tables and appropriate text provided to support recommendations, where applicable?
- Are recommendations and key clinical points displayed visually, when possible?
- Are the recommendations consistent with other ACC/AHA guidelines and other documents on the same or related topics?
Appendix Four
Introduction

Clinical Efficacy Assessment Subcommittee’s Mission Statement
The mission of the Clinical Efficacy Assessment Subcommittee is to oversee the development of guidelines that will improve the practice of medicine. The Subcommittee will provide advice, assistance, and direction to the Education Committee and to the staff of... Scientific Policy in the development and implementation of College guidelines.... The subcommittee will make recommendations: regarding appropriate clinical practices to be embodied in evidence-based guidelines; guidance on the appropriate use of these guidelines; development of new methods to enhance College guideline application to clinical practice; and identification of technology assessment issues pertinent to the College and internal medicine.

ACP-ASIM has been producing clinical practice guidelines since 1981. Guideline development started as a three-year grant called the Clinical Efficacy Assessment Project (CEAP). The Project was and still is carried out under the aegis of the Clinical Efficacy Assessment Subcommittee (CEAS).

The College appoints CEAS committee members. These are internists with expertise in primary care, health care administration, and medical and health services research. The Subcommittee provides advice, assistance, and direction to the Education Committee and to the staff of Scientific Policy in the development and implementation of College guidelines. The Subcommittee makes recommendations regarding:

- Clinical practices to be addressed by College guidelines;
- Guidance on the appropriate use of these guidelines;
- Development of new methods to enhance College guideline application to clinical practice; and
- Identification of technology assessment issues pertinent to the College and internal medicine.

The scope of CEAP guidelines has broadened and covers many aspects of internal medicine, such as screening for cancers and cholesterol, and preoperative cardiac evaluation.

CEAP guidelines target internists, but they can be easily adapted for use by subspecialists, family practitioners, and other primary care providers. CEAP aims to help College members provide the best health care possible, based on the best available current evidence. Guidelines are not designed to replace clinical judgment, rather they are meant to be evidence-based tools for clinical decision making.
CEAP Process

What is CEAP?

In 1981, the Clinical Efficacy Assessment Project, or CEAP, began as a three-year grant. The goals were:

1. To assemble and review the clinical literature on a specified topic;
2. To identify the best scientific papers; and
3. To analyze, reformulate, and present such information so that practitioners can readily determine the usefulness of diagnostic tests, procedures, or treatments.

The initial charge from CEAP was to evaluate medical advances. Early ACP guidelines addressed diagnostic tests and technologies. These guidelines focused on topics developed through surveys of the ACP membership and were chosen based on the interest of the internal medicine community. Because of CEAP's success, the program was given permanent status at ACP.

Today, CEAP is administered by the Clinical Efficacy Assessment Subcommittee (CEAS) and by the staff of the Scientific Policy Department of the ACP-ASIM.

How are ACP-ASIM Guidelines developed?

If you thought that the ACP-ASIM Guidelines were produced "by a bunch of old guys sitting around a board room," you're not alone in this misperception. ACP-ASIM guidelines are produced through a team effort involving the Scientific Policy staff, the steering committee known as the Clinical Efficacy Assessment Subcommittee (CEAS), and expert scientific collaborators.

How are topics chosen?

Choosing a topic for a clinical practice guideline is the first step in the CEAP process. Evidence reports commissioned by the Agency for Health Care Policy and Research (AHCPR) and generated by Evidence-based Practice Centers are the basis of our guidelines. These comprehensive evidence reports are systematic literature reviews and are available to the public.

Data gathering

Systematic literature reviews follow a strict protocol, and each article is assessed using standardized scoring techniques. Randomized clinical trials are given the highest evidential weight followed by prospective and retrospective cohort studies and case-control studies. The weighting is assigned based on the type of study and its methodological strength. Gaps in knowledge are also identified.
The evidence is summarized as evidence tables, the data is analyzed, and a report synthesizing the evidence is written. This evidence report, along with consensus statements where there are gaps in the evidence, becomes the foundation for the College's background paper and evidence-based clinical practice guideline.

**Review**

Both the background paper and the guideline undergo a lengthy and thorough review process. These papers are first reviewed by CEAS, and then outside reviewers are invited to comment on the manuscripts. After CEAS approves the papers, they are sent to the committee that oversees CEAS, the Education Committee.

Once the Education Committee approves the documents, they go for final approval as ACP-ASIM policy to the Board of Regents, the highest body of ACP-ASIM. Simultaneously, the Board of Governors, who represent members from all 50 states and territories as well as our international members, reviews the manuscripts.

The goal is to provide clinicians with a clinical practice guideline based on the best evidence available; to make recommendations based on that evidence; to inform clinicians of when there is no evidence; and finally, to help them deliver the best health care possible.
Appendix Five
I. BACKGROUND AND DEFINITION

In 1991, the American Psychiatric Association (APA), through its Assembly and Board of Trustees, embarked on the process of developing practice guidelines. Since its inception, the APA has generated, under many different formats, guidelines for psychiatric practice. "Practice guidelines" as defined by this project, however, are systematically developed documents appearing in a standardized format presenting patient care strategies to assist psychiatrists in clinical decision-making. Importantly, while guidelines may be used for a variety of purposes, their primary purpose is to assist psychiatrists in their care of patients.

Both the American Medical Association (AMA) and the Institute of Medicine (IOM) have sought to define the key features necessary to ensure that practice guidelines are of high quality. The AMA's attributes apply to the development process, stating that practice parameters/guidelines should: 1) be developed by or in conjunction with physician organizations, 2) explicitly describe the methodology and process used in their development, 3) assist practitioner and patient decisions about appropriate health care for specific clinical circumstances, 4) be based on current professional knowledge and reviewed and revised at regular intervals, and 5) be widely disseminated. The IOM's attributes are criteria for evaluating the finished product: validity, based on the strength of the evidence, expert judgement, and estimates of health and cost outcomes compared with alternative practices; reliability/ reproducibility; clinical applicability and flexibility; clarity; attention to multidisciplinary concerns; timely updates; and documentation.

Taken together, these prescriptives have essentially set national standards for guideline efforts.

II. TOPIC SELECTION
The APA Steering Committee on Practice Guidelines oversees the development of APA guidelines. The Steering Committee selects topics for practice guidelines according to the following criteria:

a) degree of public importance (prevalence and seriousness)
b) relevance to psychiatric practice
c) availability of information and relevant data
d) availability of work already done that would be useful in the development of a practice guideline
e) an area in which increased psychiatric attention and involvement would be helpful for the field

III. CONTRIBUTORS

APA practice guidelines are developed by a work group of psychiatrists in active clinical practice, including academicians or researchers who spend a significant percentage of their time in the clinical care of patients. Work group members are selected on the basis of their knowledge and experience in the topic area, their commitment to the integrity of the guideline development process as outlined by the AMA and IOM, and their representativeness of the diversity of American psychiatry.

Work group members are asked to decline participation if they feel there are possible conflicts of interest or biases that could impact their ability to maintain scientific objectivity. The following statement appears in every practice guideline to clarify this point:

Psychiatrists who are in active clinical practice have developed this practice guideline. In addition, some contributors are primarily involved in research or other academic endeavors. It is possible that through such activities, many contributors have received income related to treatments discussed in this Guideline. A number of mechanisms are in place to minimize the potential for producing biased recommendations due to...
conflicts of interest. Members of the APA as well as representatives from related fields have extensively reviewed the Guideline. Contributors and reviewers have all been asked to base their recommendations on an objective evaluation of the available evidence. Any contributor or reviewer who believes that he or she has a conflict of interest that may bias (or appear to bias) his or her work has been asked to notify the APA Office of Quality Improvement and Psychiatric Services. This potential bias is then discussed with the work group chair and the chair of the Steering Committee on Practice Guidelines. Further action depends on the assessment of the potential bias.

The APA is listed as the “author” of practice guidelines, with individual contributions, and reviewers acknowledged. Final editorial responsibility for practice guidelines rests with the Steering Committee and the Office of Quality Improvement and Psychiatric Services.

IV. EVIDENCE BASE

The evidence base for practice guidelines is derived from two sources--research studies and clinical consensus. Where gaps exist in the research data, evidence is derived from clinical consensus, obtained through extensive review of multiple drafts of each guideline (see section VI). Both research data and clinical consensus vary in their validity and reliability for different clinical situations; guidelines state explicitly the nature of the supporting evidence for specific recommendations so that readers can make their own judgements regarding the utility of the recommendations. The following coding system is used for this purpose:

[A] Randomized clinical trial. A study of an intervention in which subjects are prospectively followed over time; there are treatment and control groups; subjects are randomly assigned to the two groups; and both the subjects and the investigators are
"blind" to the assignments.

[B] Clinical trial. A prospective study in which an intervention is made and the results of that intervention are tracked longitudinally, that does not meet standards for a randomized clinical trial.

[C] Cohort or longitudinal study. A study in which subjects are prospectively followed over time without any specific intervention.

[D] Case-control study. A study in which a group of patients is identified in the present and information about them is pursued retrospectively or backward in time.

[E] Review with secondary data analysis. A structured analytic review of existing data, e.g., a meta-analysis or a decision analysis.

[F] Review. A qualitative review and discussion of previously published literature without a quantitative synthesis of the data.

[G] Other. Opinion-like essays, case reports, and other reports not categorized above.

The literature review process is explicitly described in the guideline, including statements concerning:

a) basic search strategy (e.g., key words, time period covered, research methodologies considered, etc.)

b) sources for identifying studies (e.g., review articles, texts, abstracting and indexing services, Index Medicus, Sciences Citations Index, computer search services, etc.)

c) criteria for selecting publications (e.g., number of relevant publications identified, whether all were reviewed, whether only prospective studies were selected, etc.)

d) review methods (e.g., publications reviewed in their entirety, abstract review only)

e) methods for cataloguing reported outcomes (e.g., study design, sample characteristics, relevant findings, etc.)
The literature review will include other guidelines addressing the same topic, when available. Wherever possible, evidence tables are constructed to illustrate the data regarding risks and benefits for each treatment. In many cases, however, evidence tables are used only to assist in writing the text and do not appear in the guideline.

V. FORMAT

Each practice guideline is presented in a standardized format, with variations as appropriate (e.g., a guideline about psychiatric evaluation or a procedure may vary from that about a specific illness).

The outline for the Major Depressive Disorders revision and subsequent guidelines and revisions is as follows:

PART A: Treatment Recommendations
I. Executive Summary of Recommendations
II. Formulation and Implementation of a Treatment Plan
III. Specific Clinical Features Influencing the Treatment Plan

PART B: Background Information and Review of Available Evidence
IV. Disease Definition, Epidemiology, Natural History
V. Review and Synthesis of Available Evidence

PART C: Future Research Needs

Reviewers and Reviewing Organizations
References

Section I provides an overview of the organization and scope of recommendations contained in subsequent sections, with each recommendation identified as falling into one of three
American Psychiatric Association
Practice Guideline Development Process
Revised 1999 Page 6

categories of endorsement:

[I] Recommended with substantial clinical confidence
[II] Recommended with moderate clinical confidence
[III] Options that may be recommended on the basis of individual circumstances

Section II presents a synthesis of the information discussed in section V, directed at providing a framework for clinical decision making for the individual patient.

Section III addresses psychiatric, general medical, and demographic factors influencing treatment, including comorbidities. Relevant ethnic, cross-cultural, social or extrinsic factors (e.g., cultural mores, family, support system, living situation, health care beliefs, etc.) which could potentially preclude or modify the practical application of guidelines and may play a role in health care decisions are emphasized.

Section IV presents the characteristics of the illness using current DSM criteria. Differential diagnosis, appropriate diagnostic procedures, aspects of the epidemiology and natural history with important treatment implications, and issues concerning special patient characteristics are outlined in this section.

Section V presents a review of the available data on all potential treatments, organized according to three broad categories: 1) psychiatric management, 2) psychosocial interventions, and 3) somatic interventions. For each treatment, this information is presented in a standard format:

a) goals of treatment
b) efficacy data
c) side effects and safety
d) implementation issues (e.g., patient selection, laboratory testing, dosing, frequency, duration)
Part C identifies directions for further research.

Immediately following the research directions is the list of individuals and organizations that submitted substantive comments of guideline drafts.

Lastly, all references cited in the published guideline are listed.

VI. REVIEW, DISSEMINATION, AND UPDATES

Each practice guideline is extensively reviewed at multiple draft stages. Draft 1 is reviewed by the Steering Committee. Approximately fifty reviewers with expertise in the topic, representatives of approximately one hundred related organizations, the APA Assembly, District Branches, Joint Reference Committee, Board of Trustees, Council on Quality Improvement, other components related to the subject area, and any APA member by request, are given the opportunity to review and comment on each practice guideline prior to publication.

The development process may be summarized as follows:

Step 1: The Steering Committee on Practice Guidelines selects a small number of individuals to serve as the work group chair and members.

Step 2: The work group chair and Office of Quality Improvement staff develop a preliminary outline, to be continuously revised and refined throughout subsequent steps in the development process.

Step 3: A literature search is conducted by APA and/or the work group. Relevant articles from the search are obtained, in abstract or in entirety. The work group reviews these articles, codes them for study design, and constructs evidence tables for each treatment.

Step 4: Draft 1 is written based upon evidence tables and
Step 5: Draft 1 is circulated to the work group and Steering Committee for review and comment.

Step 6: Draft 2 is written based on comments received.

Step 7: Draft 2 is circulated to the work group, the Steering Committee, and approximately 50 reviewers with expertise in the subject area.

Step 8: Draft 3 is written based on comments received.

Step 9: Draft 3 is circulated to the work group, Steering Committee, 50 expert reviewers, the Board of Trustees, Assembly, Joint Reference Committee, Council on Quality Improvement, Council Chairs, Commission on Psychotherapy by Psychiatrists, Committee on Women, District Branches, individual members (open review available through District Branches), the American Journal of Psychiatry and 100 representatives of related organizations.

Step 10: Draft 4 is written based on comments received.

Step 11: Draft 4 is submitted to the formal APA review and approval process (Council on Quality Improvement, Assembly, Board of Trustees).

After development of draft 4, and prior to final approval, work group members develop continuing medical education (CME) questions based on information contained in the guideline. Questions will be used for PRITE exams, ABPN recertification exams, on-line CME, and in the practice guideline monograph published by the American Psychiatric Publishing Group, Inc.

At the same time, individuals are identified to develop a corresponding quick reference guide and patient and family guide. Such individuals may include APA staff, work group
members, and/or independent contractors. These items need only to meet the approval of the SCPG as the information contained in them comes from the board-approved practice guideline.

After final approval by the Assembly and Board, each practice guideline will be widely disseminated. Guidelines will be made available to all psychiatrists in a variety of ways, including publication in the American Journal of Psychiatry. Each practice guideline will be revised at regular intervals to reflect new knowledge in the field.
Practice Guideline Development Process: A Quick Guide

What Is a Practice Guideline?

The primary purpose of a practice guideline (PG) is to assist psychiatrists in clinical decision-making and patient care.

The AMA and the IOM have described the following attributes of good PGs:

- Explicit methodology and development process
- Systematic, meticulous review and documentation of evidence
- Clear, unambiguous recommendations accompanied by descriptions of the strength of evidence behind them
- Regular scheduled review and revision
- Wide dissemination

Overview of Steps in the Development Process

A. Formation of the Work Group
- The APA Steering Committee on Practice Guidelines (SCPG) identifies a disorder (or topic) as appropriate for a new guideline (or identifies the need for revision of a previous guideline).
- The work group (WG) chair and members are selected.
- An initial outline for the PG is established that
  - follows general format for all guidelines (table 1 and pp. 2-4),
  - incorporates unique aspects of the specific topic of the PG.
- Specific topic areas are assigned to individual WG members.

B. Review of Evidence
- A systematic literature search identifies relevant published studies.
- Search results are distributed to WG members according to assigned topic.
- WG members review specific articles and generate evidence tables that
  - show key elements and findings of each study,
  - serve as a focal point for discussion and development of treatment recommendations.

C. Draft Creation and Review
- An initial draft is written that is based upon
  - review and synthesis of available evidence,
  - consensus recommendations where evidence is lacking on clinically important issues.
Multiple stages of review are undertaken, and reviewer comments are incorporated into each successive draft (table 2).

The PG is approved by the APA Assembly and Board of Trustees.

**D. Dissemination**

- Dissemination of the PG to practitioners is achieved through multiple mechanisms including publication in the *American Journal of Psychiatry*.
- Additional products are derived from the PG, including a quick reference guide and a continuing medical education program.

Throughout the development process, consideration must be given to neutralize or balance against potential sources of bias or conflict of interest.

**Step-by-Step Process to Create Draft 1**

1. **A systematic literature search identifies relevant published studies.**
   - APA staff perform an initial PubMed search of relevant MeSH subject headings.
   - More specific, subsequent searches are performed on an as-needed basis using specialized databases and/or additional key words.
   - The time period of interest will vary (searches are more extensive for new guidelines than for revisions).
   - If WG members do any searches independently, APA project staff should be given information on the search parameters and output.

2. **The WG reviews and abstracts the evidence into “evidence tables.”**
   - Evidence tables serve as the essential building blocks of the PG.
   - Tables are constructed for all relevant prospective randomized clinical trials.
   - Tables are generally constructed for other relevant clinical trials, case-control studies, and cohort or longitudinal studies.
   - Each table summarizes key elements of a study including:
     - Description of study design
     - Categorization of study design (table 3)
     - Sample characteristics
     - Relevant findings (expressed in terms of defined outcome measures whenever possible)
     - Methodological strengths and limitations
     - Other factors that may impact on interpretation of study findings
     - Clinical implication(s) of the study

p. 2 of 7
3. The WG discusses evidence tables at regularly scheduled conference calls.

The purpose of discussion is as follows:

- Review evidence tables.
  - Evidence tables are presented by individual WG members according to general subtopic.
  - Discussion should address the specifics of each study as summarized in the tables.
- Draw conclusions about the clinical implications of the aggregate evidence on a particular subtopic.
- Arrive at consensus recommendations on related issues that are clinically important but for which evidence is lacking.

As stressed by AMA and IOM, it is critical that the evidence review drive the writing of the guideline recommendations, not the other way around. Structured discussions of the evidence by the WG in conference calls will naturally lead to the next step of the development process, writing an initial draft of the guideline.

4. WG members write “summary statements” for their assigned PG sections.

These statements summarize the evidence and the recommendations agreed upon. Combined, the statements form what will become Part B, Section V, of the guideline, “Review and Synthesis of Available Evidence” (table 1).

Standard Sections of the PG, in the Order They Are Written

1. Review and Synthesis of Available Evidence

- Constitutes Part B, Section V, of the PG, but is the first section of the PG to be developed. Conclusions in this section serve as the basis for the recommendations of Part A, Section I, of the final PG.
- Explicitly states the nature of the evidence supporting specific recommendations.
- Permits readers to make their own judgments about recommendations.
- Organized according to three broad categories:
  - Psychiatric management
  - Psychosocial interventions (e.g., individual, family, and group psychotherapies; cognitive, behavioral, and psychodynamic psychotherapies, psychoeducation, psychosocial interventions)
  - Somatic interventions (e.g., pharmacotherapy, ECT)
- For each potential treatment, information is specifically presented on
  - goals of treatment;
  - efficacy data;
• side effects and safety; and
• implementation issues (e.g., patient selection; laboratory testing; treatment dosing, frequency, and duration).

For each aspect of treatment, this section includes
• initial sentence that provides an overview of the available studies,
• several sentence summaries of each of the methodologically rigorous studies (generated from evidence tables).

A concluding sentence/paragraph
• describes the overall clinical implications of the evidence,
• serves as the basis for an evidence-based treatment recommendation for Part A, Section I.

2. Formulation and Implementation of a Treatment Plan
• Constitutes Part A, Section II, of the PG.
• Addresses general features of psychiatric management such as the following:
  • Performing a diagnostic evaluation
  • Evaluating the safety of the patient and others
  • Determining a treatment setting
  • Establishing and maintaining a therapeutic alliance
  • Monitoring treatment response
  • Providing education to the patient and to the family
  • Enhancing adherence to treatment
  • Evaluating and managing functional impairments
• Addresses specific features of management such as the following:
  • Goals of treatment
  • Choice of initial treatment modality
  • Approaches for patients who do not respond to treatment initially
• Formal evidence tables are not needed but appropriate references to the literature should be included and categorized (table 3).

3. Specific Clinical Features Influencing the Treatment Plan
• Constitutes Part A. Section III. of the PG
• Addresses specific factors that may modify guideline application including:
  • Psychiatric factors (e.g. suicidality, violence or specific psychiatric symptoms or comorbidities)
  • General medical factors
  • Demographic factors influencing treatment (e.g. age, gender)
  • Ethnic and cross-cultural factors
  • Social or extrinsic factors (e.g., cultural mores, family, support system, living situation, health care beliefs, etc.)
• Formal evidence tables are not needed but appropriate references to the literature should be included and categorized (see table 3).

4. Disease Definition, Epidemiology, Natural History
• Constitutes Part B, Section IV, of the PG.
• Provides clinically relevant background information about the disorder.
• Includes brief discussion of the following:
  o Characteristics of the illness using current DSM criteria
  o Differential diagnosis and appropriate diagnostic procedures
  o Epidemiology and natural history
  o Issues concerning special patient characteristics
• Formal evidence tables are not needed for this section, but appropriate references to the literature should be included and categorized (see table 3).

5. Executive Summary of Recommendations
• Constitutes Part A, Section 1, of the PG.
• Usually written for draft 2 or 3.
• Includes a very brief summary of the key points of psychiatric management
• Consists primarily of the key treatment recommendations noted in the review of the available evidence.
• Rates the strength of each recommendation (table 4)

6. Future Research Needs
• Constitutes PART C of the PG.
• Usually written for draft 4.
• Delineates key clinical questions for which evidence is lacking (and that will have been identified throughout the PG development process).
Table 1

General Outline for Practice Guidelines

<table>
<thead>
<tr>
<th>PART A</th>
<th>Treatment Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>Executive Summary of Recommendations</td>
</tr>
<tr>
<td>II.</td>
<td>Formulation and Implementation of a Treatment Plan</td>
</tr>
<tr>
<td>III.</td>
<td>Specific Clinical Features Influencing the Treatment Plan</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PART B</th>
<th>Background Information and Review of Available Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV.</td>
<td>Disease Definition, Epidemiology, Natural History</td>
</tr>
<tr>
<td>V.</td>
<td>Review and Synthesis of Available Evidence</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PART C</th>
<th>Future Research Needs</th>
</tr>
</thead>
</table>

List of Reviewers and Reviewing Organizations

References

Table 2

Practice Guideline Draft Review Process

<table>
<thead>
<tr>
<th>Draft</th>
<th>Reviewers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>WG, SCPG</td>
</tr>
<tr>
<td>2</td>
<td>WG, SCPG, ~50 reviewers with expertise in the subject area</td>
</tr>
<tr>
<td>3</td>
<td>WG, SCPG, ~50 expert reviewers, ~100 representatives of related organizations, the APA Board of Trustees, Assembly, Joint Reference Committee, Council on Quality Improvement, Council Chairs, Commission on Psychotherapy by Psychiatrists, Committee on Women, District Branch Executives, individual APA members (open review available through District Branches), and the American Journal of Psychiatry.</td>
</tr>
<tr>
<td>4</td>
<td>Final review through the formal APA review and approval process (i.e., Council on Quality Improvement, Assembly, Joint Reference Committee, Board of Trustees)</td>
</tr>
</tbody>
</table>
Table 3

<table>
<thead>
<tr>
<th>Categorization of Study Designs and Nature of Supporting Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Randomized clinical trial</td>
</tr>
<tr>
<td>A study of an intervention in which subjects are prospectively followed over time; there are treatment and control groups; subjects are randomly assigned to the two groups; and both the subjects and the investigators are &quot;blind&quot; to the assignments.</td>
</tr>
<tr>
<td>B Clinical trial</td>
</tr>
<tr>
<td>A prospective study in which an intervention is made and the results of that intervention are tracked longitudinally, that does not meet standards for a randomized clinical trial.</td>
</tr>
<tr>
<td>C Cohort or longitudinal study</td>
</tr>
<tr>
<td>A study in which subjects are prospectively followed over time without any specific intervention.</td>
</tr>
<tr>
<td>D Case-control study</td>
</tr>
<tr>
<td>A study in which a group of patients is identified in the present and information about them is pursued retrospectively or backward in time.</td>
</tr>
<tr>
<td>E Review with secondary data analysis</td>
</tr>
<tr>
<td>A structured analytic review of existing data, e.g., a meta-analysis or a decision analysis.</td>
</tr>
<tr>
<td>F Review</td>
</tr>
<tr>
<td>A qualitative review and discussion of previously published literature without a quantitative synthesis of the data.</td>
</tr>
<tr>
<td>G Other</td>
</tr>
<tr>
<td>Opinion-like essays, case reports, and other reports not categorized above.</td>
</tr>
</tbody>
</table>

Table 4

Ratings of the Strength of Treatment Recommendations

[I] Recommended with substantial clinical confidence

[II] Recommended with moderate clinical confidence

[III] Options that may be recommended on the basis of individual circumstances
Appendix Six
THE DEVELOPMENT OF EVIDENCE-BASED CLINICAL PRACTICE GUIDELINES

Integrating Medical Science and Practice

Richard T. Connis
David G. Nickinovich
American Society of Anesthesiologists

Robert A. Caplan
University of Washington

James F. Arens
M.D. Anderson Cancer Center

Abstract

Practice guidelines are rapidly becoming preferred decision-making resources in medicine, as advances in technology and pharmaceutics continue to expand. An evidence-based approach to the development of practice guidelines serves to anchor healthcare policy to scientific documentation, and in conjunction with practitioner opinion can provide a powerful and practical clinical tool. Three sources of information are essential to an evidence-based approach: a) an exhaustive literature synthesis; b) meta-analysis; and c) consensus opinion. The systematic merging of evidence from these sources offers healthcare providers a scientifically supportable document that is flexible enough to deal with clinically complex problems. Evidence-based practice guidelines, in conjunction with practice standards and practice advisories, are invaluable resources for clinical decision making. The judicious use of these documents by practitioners will serve to improve the efficiency and safety of health care well.

Keywords: Practice guidelines, Evidence-based medicine

Practice guidelines in medicine have traditionally been created as policy documents that serve as information resources for the systematization of clinical practice. Their intended purpose is to provide physicians and other healthcare professionals with a useful reference for optimizing patient care. Because guidelines are usually developed and endorsed by a healthcare organization with the intent of regulating or standardizing clinical decision making (4-6), the broadest possible base of evidence is critical to their development. A broad evidence base will include a comprehensive assessment of peer-reviewed scientific literature combined with interpretations based on the clinical experience of practitioners (27).
Historically, the use of scientific literature in the development of practice guidelines has been selective rather than systematic. Prior to the 1980s, literature reviews were typically narrative, with the search process driven by the subjective judgment of reviewers (20;25). This “traditional” approach to literature review is often limited by the reviewers’ knowledge of the literature and inclusion of a disproportionate number of articles supportive of the reviewers’ viewpoints (8;18). According to Chalmers and Lau (3).

Too often, authors of traditional review articles decide what they would like to establish as the truth either before starting the review process or after reading a few persuasive articles. They then proceed to defend their conclusions by citing all the evidence they can find. The opportunity for a biased presentation is enormous, and its readers are vulnerable because they have no opportunity to examine the possibilities of biases in the review.

Evidence-based approaches have the potential to avoid systematic bias through the combination of a structured, exhaustive evaluation of scientific documentation and an assessment of diverse practitioner opinion (5). The application of quantitative techniques and precise rules to combine research findings from various independent studies bolsters the scientific rigor of the aggregated literature with meta-analysis as the primary approach. However, it is important to note that meta-analytic results alone can be misinterpreted as easily as the results of an individual study. An evaluation of scientific documentation in the appropriate clinical context is aided by surveys and other documented opinions from experts and practicing healthcare providers.

In applying meta-analytic and other scientific findings, the information provided by the practice guideline in the form of recommendations needs to be flexible enough to accommodate the complexities of clinical practice. Analytic evidence may indicate that a treatment or other intervention is effective. However, a guideline recommendation needs to allow for the clinical judgment of the practitioner, who determines whether the intervention is medically warranted or appropriate for a specific case. An additional source of evidence is needed in developing a guideline, and is best obtained by evaluating information based on the clinical experiences of experts and practitioners. Scientific knowledge can then be meaningfully combined with clinical judgment to develop recommendations for the application of a designated intervention (28). A guideline must also be feasible for use in a wide range of practice settings, meaning that scientific evidence and expert opinion should be supplemented by opinions from the broader population of practitioners (27). This broad base of opinion can benefit the development and implementation of a guideline in two ways. First, input from a variety of practice settings (e.g., large academic institutions and small rural settings) may contribute to a guideline’s flexibility by identifying issues and problems unique to each distinct setting. Second, constructive forums for the expression of divergent opinions prior to a guideline’s formal implementation enhance a guideline’s acceptance by the general membership of a medical specialty.

The purpose of this paper is to define and describe elements of a multifaceted guideline development process currently used by the professional association of one medical specialty, the American Society of Anesthesiologists (ASA). The ASA has published 10 evidence-based practice guidelines, including guidelines for difficult airway management, acute pain, chronic pain, cancer pain, preoperative fasting, and obstetrical anesthesia (2;7;9;10;13;14;19;21;23;24). These guidelines have been well received.

Practice guidelines were developed by the ASA to address issues that could not be reasonably addressed by practice standards. Practice standards typically provide specific requirements for practice and are applied, with few exceptions, to virtually all relevant clinical situations. Although standards are important prescriptions for anesthesia care, it was recognized that more complex topics warranted a less rigid approach. Practice guidelines
were recognized as tools for providing clinical recommendations that would address these broader topics. Because of the complexity of the issues addressed, it was necessary for guidelines to incorporate a comprehensive array of evidence, including detailed assessments of the scientific literature and consensus documentation from multiple sources.

IDENTIFICATION OF GOALS AND OBJECTIVES

ASA's evidence-based guidelines generally take 1–3 years to develop, and include a number of important steps before final completion and approval (Table 1). Typically a policy committee or task force is convened, consisting of academic and clinical practitioners recognized as experts in the topic of concern and representing a variety of practices and

Table 1. Protocol for Practice Guideline Development

1. Committee/task force assignment
2. Identification of potential evidence linkages
3. Literature search
   a. Articles considered (original studies or reports published in peer-reviewed journals)
      1) Randomized controlled trials
      2) Nonrandomized comparative studies
      3) Controlled observational studies
      4) Retrospective comparative studies
      5) Uncontrolled observational studies
      6) Case reports
   b. Articles not considered
      1) Letters with no original data
      2) Editorials, review articles, and commentaries
      3) Meta-analytic studies (these analyses use data generated from other studies)
      4) Personal correspondence
      5) Unpublished papers/presentations
4. Availability of scientific evidence in the literature is determined. If none of the evidence linkages has sufficient literature at this point in the process, a decision is made to either revise the evidence linkages or produce a practice advisory.
5. Literature synthesis with assessment of directional evidence
   a. Review and sort studies into potential evidence linkage categories
      1) Record relevant information related to clinical factors (e.g., patient health status, clinical interventions used, health outcomes).
      2) Code information related to statistical evidence (e.g., study design, statistical tests, significance levels)
   b. Assign directional support for a potential evidence linkage addressed by each selected study (some studies address multiple linkages). For each study, determine direction related to patient benefit (positive, negative, or neutral).
   c. Determine overall direction of support for evidence linkage by summation of individual studies.
6. Hypothesis development
   a. Assess overall linkage directions.
   b. Determine one-tailed hypotheses based on linkage direction.
7. Meta-analysis: Adequately designed studies with sufficient quantitative information to describe a statistical relationship between a clinical intervention and a clinical outcome are identified.
   a. Randomized controlled trials
   b. Nonrandomized comparative studies (conditionally acceptable)
8. Consensus assessment
   a. Surveys related to evidence linkages
      1) Expert consultants
      2) Broad representation of practitioners
   b. Feasibility surveys
   c. Open forum presentations
   d. Internet commentary
9. Formal review and approval by specialty organization
Connis et al.

geographic areas. This broad representation leads to improved chances of generalizing the final recommendations.

The task force begins the process by formalizing the intended topics, goals, and objectives for the proposed guideline. These items include a description of the clinical disorders and conditions to be addressed by the guidelines, the types of patients for whom the guidelines are intended, clinical interventions (e.g., diagnostic tests, treatments) that will be considered in developing the guidelines, the principal intended users of the guidelines, and the practice settings in which the guidelines are applicable. These formalized topics play a crucial role in defining the scope of the guideline. Once the formalized topics, goals, and objectives are identified and agreed upon by the task force, an evidence model is formulated. The evidence model specifies criteria for inclusion/exclusion of data from the literature or other sources.

**DEVELOPING AN EVIDENCE MODEL**

Following specification of the guideline's goals and objectives, a series of potential evidence linkages is formulated (26). Evidence linkages represent statements about relationships between clinical interventions and clinical outcomes. A clinical intervention is typically an activity performed by a physician or other healthcare provider (e.g., administering a specified drug). A clinical outcome is recorded in terms of its potential benefit to the patient (e.g., reduced pain or minimization of specified side effects). An important component of an evidence linkage is the specificity of the identified interventions and outcomes. For example, “analgesics provide maternal pain relief” would be further specified with a statement such as “epidural bupivacaine with opioids affects maternal analgesia during labor.” This specification of targeted interventions and outcomes, in conjunction with other elements from the evidence model (e.g., intended providers, targeted patient population), will provide definitive direction for obtaining, organizing, and evaluating the evidence.

This evidence model provides the framework for a guideline's clinical recommendations, and essentially provides the structure for the entire development process. Once the model is in place, the task force can initiate a multistep process. The elements of this process will include literature searches, literature syntheses, meta-analyses, survey development, consensus evaluation, feasibility studies, open-forum presentations, Internet commentary, and formal endorsement by the society or healthcare organization.

**LITERATURE SEARCH FOR EVIDENCE**

The literature search usually includes a computerized search of large reference sources, such as the National Library of Medicine or Nursing and Allied Health Abstracts. Other electronic resources, also readily available on CD-ROM and/or the Internet, contain databases of reviews and abstracts as well as full-text articles. Software for bibliographic reference databases is an important tool in the search process, citation management, and dissemination of findings. Manual searches of literature are conducted to supplement electronic sources. Since electronic searches are typically driven by keyword search engines, they do not always have appropriate search mechanisms to locate relevant topics of interest.

To be useful for evidence-based guideline development, studies must meet certain criteria. First, a study must report a clinical finding or set of findings that can be tallied or quantified. This requirement eliminates reports that contain only commentary or undocumented opinions of the authors. Second, a study must be an original investigation or report containing a clinical finding or set of findings. Thus, review articles or manuscripts that report findings from other sources are not used. For meta-analytic evidence, study findings
Evidence-based guideline development must clearly indicate a specified relationship between a clinical intervention and an outcome of interest.

After the first phase of article search and classification is completed, a listing of all articles located to date is presented to the task force. Members are asked to review all articles listed, both those accepted and those not accepted, and to suggest changes in the acceptance or nonacceptance of individual articles when needed. They also may add articles not listed, to refine and expand the search process.

Some practice guidelines developed by other organizations may include unpublished literature as potential evidence assessment. Use of such literature addresses the issue of sampling or "publication" bias, in which journals may be biased toward accepting articles that report statistically significant findings (25). As a result of such bias, many manuscripts reporting nonsignificant findings or no differences between groups or conditions are not accepted for publication. Often, upon discovering that they have no significant findings, investigators may not bother to submit a manuscript for publication (i.e., a "file-drawer" problem) (15).

Although publication bias is a viable concern, ASA investigators do not currently use unpublished literature as a source of data for several reasons. Unpublished studies are not peer-reviewed and the use of such literature may incorporate data derived from inadequate research methods, the use of inappropriate or flawed statistical procedures, and other problems usually identified and corrected by the peer-review process. In addition, investigators do not know whether the obtained sample of unpublished studies is representative of the relevant population of unpublished literature. Therefore, this potential for selection bias is as serious a threat as that of publication bias. Although some research groups have endeavored to create repositories of unpublished studies, the completeness of such databases remains questionable. In lieu of obtaining a representative sample of unpublished studies, standard statistical methods (e.g., computation of a "fail-safe" N value) are available and provide a reasonable estimate of the required number of additional (i.e., unpublished) studies reporting contradictory outcomes sufficient to nullify the findings obtained from published studies (16;25).

LITERATURE SYNTHESIS WITH DIRECTIONAL ASSESSMENT

Evidence linkages are initially used for purposes of identifying relevant literature. This literature is reviewed and detailed information is extracted, including but not limited to patient data (e.g., clinical condition, age); data regarding the treatment, procedure, or anesthetic intervention: outcomes reported; and research design and statistical analyses. Spreadsheet technology applied to this task is invaluable, particularly in the subsequent management of the data and summarization of findings. The use of such technology can greatly reduce the time and effort spent aggregating the data.

In their spreadsheet documentation, the ASA includes a classification value of the predominant "direction" of study findings. For each reviewed study, the outcome of interest is classified as supporting a linkage, refuting a linkage, or neutral. Each article is coded (i.e., support = +1, refutation = −1, neutral = 0), and a summary value is calculated across all studies. From these results, a directional (one-tailed) assessment of support or refutation for each linkage is obtained. Following the directional assessment, the evidence linkages are revised to include directionality, therefore providing justification for the use of one-tailed statistical testing. A directional statement derived from the evidence linkage example cited earlier would be: "epidural bupivacaine with opioids improves maternal analgesia during labor compared to equal concentrations of epidural bupivacaine without opioids." All studies with data, regardless of methodology, are included in the directional assessment. No attempt is made to calculate average values or other aggregate statistics.
Connis et al.

The directional overview of the literature is viewed as a tool to refine the evidence linkage in order to initiate statistical procedures (i.e., meta-analyses). The directional assessment represents a separate and vital component of the literature-based evidence because it is an examination of all studies, including those for which effect size estimates are not provided (i.e., case reports, descriptive studies, correlational studies). These studies are important to evaluate because they contain information not necessarily found elsewhere, and their inclusion can affect the directional assessment. For example, case reports may provide an indication of adverse outcomes or previously unrecognized benefits not recorded in the clinical trial literature. Moreover, descriptive studies provide evidence related to the frequency of occurrence of an adverse or beneficial outcome when an intervention of interest is employed.

ANALYTICAL PREPARATION

Once the directional overview is complete, the evidence linkages are refined to include one-tailed hypotheses so that formal meta-analyses can proceed. At this point, only controlled comparative studies are considered for analysis. Controlled studies provide a vital indication of the effectiveness of a medical intervention.

In the meta-analytic procedures utilized by the ASA, the primary interest has been to combine original (primary) research studies for purposes of investigating questions of causality. In this effort to document causal relationships, the design features of the various studies under review are of critical importance. If the studies comprising the primary research literature have design features that assure a high level of internal validity (e.g., random assignment of subjects to conditions of the study, researcher and practitioner blinding, and researcher control over the intervention), then questions regarding causality can be addressed. When threats to internal validity are evident in the primary studies under review (e.g., treatments administered to pre-existing groups), then analyses cannot directly address questions of causality and are limited to questions of covariation. In summary, literature review can generally determine associations between variables of interest, but the investigation of questions of causality is critically dependent on the inferential robustness of the controlled studies under review.

ANALYTICAL MODEL

Meta-analysis

When an evidence linkage contains a sufficient number of studies (e.g., five or more) with well-defined experimental designs and statistical information, formal meta-analyses are conducted. A fixed-effects model using odds ratios or combined probability tests is applied when there is an expectation of minimal variation in effect size estimates among the studies in the analysis.

A fixed-effects model is used more commonly for several reasons. Historically, the anesthesia literature has used the same or very similar outcome measures (e.g., visual analog scale scores for pain measurement). Variability in outcomes among the various independent studies has generally been homogeneous. Conceptually, directional nonrandom outcomes are generally expected in the anesthesia literature. On occasion, a random-effects model may be considered when appreciable effect size variability is expected.

Usually, more than one test statistic is obtained in a meta-analysis related to a particular evidence linkage. For a meta-analysis to be supportive of an evidence linkage, all component analyses must be in agreement regarding effect sizes and significance values. As a further assurance of the congruity and robustness of the findings, all meta-analyses
Evidence-based guideline development should be in agreement with the directional assessment as well as with consensus opinion. When agreement is not apparent in any of the specified areas of evidence, further evaluation is necessary, and, if the disagreement continues to persist, the discrepancy is reported and discussed in the guideline.

Methodologic controls

As a methodologic control for reviewer bias, additional assessments of the reviewed literature are conducted independently by the task force members and methodologists. The ASA uses a sample of reviewed articles randomly selected from each evidence linkage and a random sample of articles not accepted into the database to assess agreement for study design, type of statistical analysis, identification of evidence linkage, and the reviewers’ judgment as to whether the study should be included in the database. Interobserver agreement among task force members and methodologists is assessed and reported using agreement levels for two-rater agreement pairs (17) and for multi-rater chance-corrected agreement (11;12).

Following review of the literature, tests for heterogeneity of findings from the independent studies are conducted to ensure consistency among the study results. To control for potential publishing bias, the ASA calculates a fail-safe N value for each combined probability test. A fail-safe N refers to the number of additional studies necessary to increase the overall probability value obtained to a value higher than the critical value for statistical significance (16). To ensure that the literature considered is peer-reviewed, no search for unpublished studies is conducted. The ASA does not conduct reliability tests for locating research results, because their intent is to obtain an entire population of published studies for each evidence linkage rather than collecting a representative sample of studies.

CONSENSUS AS EVIDENCE

Research findings from published literature provide the cornerstone for guideline recommendations. However, published studies alone may not provide necessary or complete information regarding relevant details of clinical practice. Accordingly, additional sources of information and evidence are actively and deliberately sought by the ASA. Such information may best be obtained from clinical experience. For example, studies examining preoperative testing may provide information about the sensitivity and specificity of a particular test without providing insight about when or on whom a test should be performed. Practitioner opinions may serve the role as a supplemental source of evidence, reflecting current practice. Topics that are addressed by obtaining practitioner opinion include issues related to the importance and practicality of the interventions identified in a guideline, and issues related to the projected cost, estimated practice time, and feasibility of implementing a guideline. Practitioner opinion may be obtained through several mechanisms, ranging from the simple recording of consensus within a designated task force to large-scale surveys and feedback from presentations or open forums at national conventions.

The ASA obtains consensus data from multiple sources, including surveys of expert consultants and of the broader population of practitioners, and open forum presentations, Internet commentary, and feasibility studies. Expert consultants are carefully chosen to provide a balance between private practice and academia, as well as representation from each of the major geographic areas of the United States. Consultants are asked to participate in surveys of their opinions of various aspects of a guideline and to review and comment on initial draft reports. Random samples of the ASA membership are also surveyed regarding the topics addressed by the evidence linkages.

Each task force holds one or more open forums at a major national anesthesia meeting to solicit input on its draft guideline from meeting attendees. During each open forum, audience
Connis et al.

testimony is recorded. Directly following an open forum, the task force meets, commentary is discussed, and clarifications in the draft document may be made. Major issues, when they arise, may require a new literature synthesis or additional consensus surveys. The revised draft is then disseminated to various additional sources (e.g., the Internet, ASA district directors, presidents of ASA component societies for their commentary. The consultants are surveyed one final time to assess their opinions on the feasibility of implementing the revised guideline recommendations. All available information is used by the task force to produce a final document for submission to the ASA for formal approval.

The term consensus in this sense refers to an evaluation of the combined agreement derived from the sampling of opinions from scientists, experts, academicians, and clinicians. Surveys used to elicit and measure these opinions are designed so that responses are easily interpreted and differences, when they occur, are clearly noted. The opinions of survey respondents are based on sets of items that are deliberately the same for each group of respondents. Responses from several sources can thus be conveniently compared.

Each consensus survey item is derived from a specific evidence linkage. Since the evidence-linkages are also employed in the literature search and assessment procedures, the same set of intervention/outcome relationships provides a constant foundation throughout the entire guideline development process.

The use of consensus as a source of evidence has not been thoroughly explored. However, consensus data often provide critical feedback on the feasibility of the proposed recommendations. “Consensus as evidence” is a relatively distinct component of ASA’s process. These data are analyzed in the same manner as directional evidence, and thus become the third major evidence source for the formulation of viable recommendations.

GUIDELINE RECOMMENDATIONS

Guideline recommendations are based directly on the evidence linkages developed at the beginning of the process. Each of the three sources of evidence (i.e., directional, meta-analytic, and consensus) is separately considered in the formulation of the final recommendations.

Agreement among the three sources of evidence is required for the full support of a recommendation related to a particular evidence linkage. Occasionally, divergence from the full support of all three sources may occur. For example, meta-analytic and directional results may support a designated intervention, but their application in clinical practice may be questioned by consensus findings. These divergences are noted, then discussed and interpreted in the guideline. By discussing the strengths and shortcomings of each recommendation, a guideline becomes sufficiently dynamic to respond to the diverse requirements of clinical practice.

CONCLUSIONS

The process described herein represents the collective efforts by the ASA to produce timely and clinically relevant guidelines. In particular, the use of an exhaustive literature search, literature synthesis, directional summarization, meta-analyses, and consensus assessment all combine to produce a multifaceted overview of evidence for rational policy decision making in clinical medicine. These activities are made easier and more accurate through the use of existing technology, such as reference databases (for literature searches and citation management), spreadsheets (for literature aggregation and analysis), and statistical software (for survey assessment).

The ASA’s evidence-based model for guideline development combines literature synthesis and analysis with the knowledge and experience obtained from clinical practice. The accumulated scientific data coupled with the recognition that practitioners on occasion
Evidence-based guideline development

may need to modify procedures to fit individual cases form a compelling basis for the widespread acceptance and use of practice guidelines.

When there is not sufficient information available for the development of an evidence-based practice guideline, an alternative may be considered. One alternative that has recently been implemented by the ASA is the practice advisory (1,2). The intent of a practice advisory is to systematically use an evidence-based model without meta-analytic findings. Therefore, until meta-analytic evidence becomes available, a practice advisory, in the form of a published report, may be used as a viable reference document for clinical practice.

Through the dissemination of practice advisories in conjunction with practice standards and guidelines, the ASA makes available a complete package of advice to clinicians on selected topics. Practice standards offer guidance for narrow and well-defined areas of practice, while guidelines and advisories are intended to address the more complex aspects of patient care and may provide information on the impact of new medical technologies or other recent interventions. These three sources of guidance, when periodically updated, will offer practitioners access to the most recent collective knowledge relating to patient care.

REFERENCES


