

## GUIDELINES

# Clinical practice guideline development manual: A quality-driven approach for translating evidence into action

Richard M. Rosenfeld, MD, MPH, and Richard N. Shiffman, MD, MCIS,  
New York, NY; and New Haven, CT

*No sponsorships or competing interests have been disclosed for this article.*

### ABSTRACT

**BACKGROUND:** Guidelines translate best evidence into best practice. A well-crafted guideline promotes quality by reducing health-care variations, improving diagnostic accuracy, promoting effective therapy, and discouraging ineffective—or potentially harmful—interventions. Despite a plethora of published guidelines, methodology is often poorly defined and varies greatly within and among organizations.

**PURPOSE:** This manual describes the principles and practices used successfully by the American Academy of Otolaryngology—Head and Neck Surgery to produce quality-driven, evidence-based guidelines using efficient and transparent methodology for action-ready recommendations with multidisciplinary applicability. The development process, which allows moving from conception to completion in 12 months, emphasizes a logical sequence of key action statements supported by amplifying text, evidence profiles, and recommendation grades that link action to evidence.

**CONCLUSIONS:** As clinical practice guidelines become more prominent as a key metric of quality health care, organizations must develop efficient production strategies that balance rigor and pragmatism. Equally important, clinicians must become savvy in understanding what guidelines are—and are not—and how they are best utilized to improve care. The information in this manual should help clinicians and organizations achieve these goals.

© 2009 American Academy of Otolaryngology–Head and Neck Surgery Foundation. All rights reserved.

If you use or develop clinical practice guidelines, this manual will likely be of interest. “There are many paths to the top of the mountain,” suggests an old Chinese proverb, “but the view is always the same.”<sup>1</sup> Although many paths lead to guidelines, we offer proven strategies for crafting a valid and action-ready product within 12 months. The driving force is quality improvement with a continuous effort to balance pragmatism with developmental rigor. The end product is a starting point for performance improvement.

This manual builds on an earlier publication<sup>2</sup> by American Academy of Otolaryngology—Head and Neck Surgery (AAO-HNS) to systematize internal guideline development. By following these principles, the AAO-HNS published five multidisciplinary guidelines in five years, *all within 12 months from conception to completion*.<sup>3-7</sup> Each guideline presented a fresh opportunity to test and refine prior efforts, necessitating a revised and greatly expanded manual only three years after initial publication. Our new manual not only summarizes this experience, but allows other organizations to assess and adapt the processes.

Our goals in publishing a revised manual are several. First, we sought to provide clinicians with a straightforward explanation of guidelines, considering the increasing prominence of guidelines as a quality metric. Secondly, we wanted a pragmatic resource, which accurately reflects current practices, to sustain consistent guideline development at the AAO-HNS. Lastly, we wanted to share our successful development process with the guideline community at large to encourage an exchange of ideas and to promote best practices.

Guidelines are particularly important when wide regional variations exist in managing a condition. Similarly, the wide variability in guideline *methodology*, both within and between organizations, is precisely what mandates a systematic approach to guideline development. Despite a plethora of techniques reflected in published guidelines, we could not find a single, comprehensive “how-to” manual with a valid and pragmatic approach that could be readily implemented. This work is offered to address this void.

We thank the AAO-HNS for their trust, support, and flexibility throughout this fruitful collaboration, and sincerely hope that you may also benefit from the experience. We humbly acknowledge that ours is one of many paths to the mountain top, and look forward to further refinement based on reader feedback and ongoing experience.

Received April 14, 2009; accepted April 20, 2009.

## HOW TO USE THIS MANUAL

Throughout the manual we emphasize *principles* and *practices*, recognizing that both are needed to translate concepts into action. *Principles* underlying practices are always stated, to promote conceptual focus and clarity before getting side-tracked with implementation details. *Practices* are illustrated with examples from prior AAO-HNS guidelines to clarify how we chose to implement a principle, with the understanding that other development groups will need to modify the particulars to fit their organizational structure and resources.

Content that is offset from the remainder of the text is intended to emphasize a concept, insight, or principle of special note or importance. This also serves to create visual breaks that improve readability, along with tables, bulleted lists, numbered lists, and frequent subheadings.

The following list describes some of the fundamental principles underlying guideline development that are discussed sequentially in this manual:

1. Medicine and guidelines: why guidelines are essential for quality care
2. Understanding guidelines: what makes a guideline useful and valid?
3. Principles of guideline development: essential steps and processes
4. Identifying evidence: finding and using best published evidence
5. Identifying topics: how to prioritize quality improvement opportunities
6. Understanding key action statements: the backbone of a clear and usable guideline
7. Understanding evidence profiles: how they promote transparency
8. Understanding recommendation grades: the link between action and evidence
9. Appraising implementability: maximizing the chance to influence clinician behavior

This manual is also intended as a practical resource for use during guideline development. Major sections of the text correspond to the activities involved (Table 1), allowing the user to move from one section to the next as development proceeds. When significant principles apply to an activity they are discussed within, or just before, the relevant section.

We have tried to make this manual as reader-friendly as possible, but how to best approach the material will depend upon one's background and perspective. Individuals and

**Table 1**  
**Timetable for guideline development**

Month	Activity	Goals
0	Planning	Define topic; identify leadership, partner organizations, and working group members
0-1	Stage 1 literature search	Identify existing guidelines and systematic reviews
2	Conference call #1	Define purpose, timeline, and scope; discuss conflicts of interest; plan stage 2 literature search
2-3	Stage 2 literature search	Identify randomized controlled trials
3	Conference call #2	Refine scope and definitions; generate a draft topic list of opportunities for quality improvement
4	In-person meeting #1	Construct a "straw man" guideline of key action statements based on topic priorities; outline supporting text for key statements; discuss writing assignments
4-5	Stage 3 literature search	Identifying best evidence to facilitate writing assignments for specific action statements
4-5	Writing assignments	Write the amplifying text for key action statements; chair collates into guideline draft
6	In-person meeting #2	Refine the key action statements; review amplifying text, assign evidence profiles; grade recommendations
6-7	Writing assignments	Revise and polish the draft guideline
7	Appraising draft guideline implementability	Appraisal of draft guideline clarity, quality, and ability to be successfully implemented
8	Conference call #3	Review guideline appraisal report; remedy deficiencies
9	Pre-release peer review	External review of draft guideline by representatives of target audience and practice settings
10	Organizational board review	Internal review and approval of final guideline by the board or directors of the sponsoring organization(s)
11-12	Publication	Final guideline submitted for publication

organizations involved in guideline development comprise a diverse audience that may benefit from the material in the manual in several ways:

- *Organizations new to guideline development* will benefit from understanding the nuances and complexities of guideline creation, and may consider using the process outlined as a starting point for their own endeavors.
- *Organizations with an established guideline development process* may wish to compare their current processes to those described here, potentially identifying areas for improved quality, efficiency, or both.
- *Members of guideline panels or working groups* can derive greater insight and understanding of development methodology, allowing them to contribute most effectively as authors or participants.
- *Staff supporting guideline development* will find practical suggestions on staying focused and efficient, and may consider using, or adapting, the manual as a template for their own processes.

Guideline users are perhaps even more diverse than guideline developers. Whereas most guideline users do not need, or want, detailed information on methodology, some understanding is necessary to interpret and utilize existing products:

- *Organizations that use guidelines* will gain greater insight into the guideline development process, understanding what makes a guideline valid and action-ready, and what attributes should be sought when critically assessing a guideline's potential value.
- *Organizations that conduct systematic reviews* can interact most effectively with guideline developers if they understand the process of moving from evidence to action, how systematic reviews facilitate these efforts, and how recommendations are made when evidence is absent or of low quality.
- *Clinicians with an interest in guidelines* can use them most effectively if they understand how they are developed, what are current best practices, and how guidelines should—and should not—be used to influence clinical care.
- *Consumers or consumer groups* that have identified guidelines of interest will be better able to assess quality, and select among multiple guidelines of varying quality, if they appreciate the processes involved in creating valid guidelines.

## MEDICINE AND GUIDELINES

The Institute of Medicine has identified three crucial tasks for a national system to identify highly effective health-care services: priority setting, evidence review, and developing recommendations (guidelines).<sup>8</sup> The last task—creating

clinical practice guidelines—is perhaps the most challenging, because methodology continues to evolve, the quality and relevance of available evidence is highly variable, and evidence gaps mandate valid processes for incorporating expert consensus.

Guidelines help clinicians translate best evidence into best practice. A well-crafted guideline promotes quality by reducing health-care variations, improving diagnostic accuracy, promoting effective therapy, and discouraging ineffective—or potentially harmful—interventions.

This manual offers one approach to efficient guideline development based on experience of the American Academy of Otolaryngology—Head and Neck Surgery (AAO-HNS) and the Yale Center for Medical Informatics. The AAO-HNS and its associated Foundation sponsor continuing medical education, professional meetings, scientific research, and practice management guidance for more than 13,000 ear, nose, and throat specialists in the United States and abroad. Since 2004 the AAO-HNS has devoted substantial resources to creating and publishing guidelines that improve quality of care in diverse clinical practice environments.

Despite an immediate need for valid, action-ready guidelines, many barriers exist. Published guidelines, although numerous, are often poorly suited to assess performance or influence care, because recommendations do not translate into measurable actions or activities. Moreover, the development process is generally inefficient and highly complex, requiring, on average, about two to three years per guideline. Gaps in the evidence base for many important issues typically preclude guideline recommendations based on evidence, even when quality concerns or practice variations mandate urgent action.

One solution is to produce *quality-driven, evidence-based* guidelines using *efficient* and *transparent* methodology for *action-ready* recommendations with *multidisciplinary applicability*:

- **QUALITY-DRIVEN** means placing quality improvement at the forefront of guideline development, using current best evidence and multidisciplinary consensus to prioritize recommendations. Selection of key action statements is driven by opportunities to promote best practices, reduce variations in care, and minimize inappropriate care or resource utilization.
- **EVIDENCE-BASED** means supporting all decisions with the best available research evidence identified through systematic literature review. An absence of high-quality evidence (eg, randomized trials), however, does not preclude a structured use of expert consensus if an important quality concern needs to be addressed.

- EFFICIENT guidelines make maximum use of available resources to create a timely product, ideally moving from conception to publication within 12 to 18 months.
- TRANSPARENT METHODOLOGY is explicit, reproducible, and applied consistently so guideline users can link recommendations to the corresponding level of evidence, benefit-harm-cost relationship, and the roles of values and patient preferences in decision making.
- ACTION-READY recommendations tell providers what to do, to whom, under what specific circumstance, using unambiguous language that facilitates implementation and measurement.
- MULTIDISCIPLINARY validity and applicability means that all stakeholders (eg, primary care, specialists, allied health, nursing, consumers) are part of the development and implementation processes.

The utility of a guideline depends highly on its *transparency*, which makes clear the purpose and basis of recommendations to end users. Transparency mandates disclosure of competing interests by authors, explicit statements about the reasons for developing a policy, and explanation of contributing factors are weighed.<sup>9</sup>

AAO-HNS guidelines prescribe recommendations in *key action statements* followed by amplifying text. All guideline action statements should ideally be supported by *evidence profiles* that summarize clearly the decision-making process in terms of aggregate evidence quality, harm-benefit assessment, development group values, and the role of patient preference. Evidence profiles are discussed fully later in this manual.

## UNDERSTANDING GUIDELINES

As defined by the Institute of Medicine, clinical practice guidelines are “systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances.”<sup>10</sup> Despite increasing acceptance of an evidence-based approach to clinical decision making, much clinical practice is still not based on the best available evidence. Guidelines are one way of implementing evidence into practice.<sup>11</sup> They can serve as a guide to best practices, a framework for clinical decision making, and a benchmark for evaluating performance.

Guidelines benefit patients through better outcomes, fewer ineffective interventions, and greater consistency of care, and by creating secondary implementation materials (pamphlets, videos, etc). Clinicians can use guidelines to make better decisions, initiate quality improvement efforts, prioritize new research initiatives, and support coverage or reimbursement for appropriate services. Conversely, a

flawed guideline could significantly harm both patients and clinicians, thereby mandating sound methodology as a basis for guideline development.<sup>12</sup>

Simply inserting the word “guideline” in the title of a document does not make it so. Many review articles, consensus statements, practice parameters, and policy recommendations are mistakenly labeled as “guidelines,” even though they do not possess the methodologic rigor to warrant such a designation. A real guideline is one that fulfills all or most of the specific criteria defined below.

The Appraisal of Guidelines Research & Evaluation (AGREE) instrument is a widely used generic measure of guideline quality.<sup>13</sup> Quality guidelines are characterized by the following attributes:

1. EXPLICIT SCOPE AND PURPOSE: Specific descriptions are given of the overall guideline objective(s), the clinical question(s) covered, and the patients to whom the guideline is meant to apply.
2. STAKEHOLDER INVOLVEMENT: The development group includes individuals from all relevant professional groups; patients’ views and preferences are sought; target users are clearly defined; and the guideline has been piloted among target users.
3. RIGOR OF DEVELOPMENT: Systematic methods are used to search for and select evidence; methods for formulating recommendations are clearly described; recommendations take into account health benefits, side effects, and risks; recommendations are linked explicitly to supporting evidence; the guideline is externally reviewed by experts prior to publication; and a procedure for updating the guideline is provided.
4. CLARITY OF PRESENTATION: Recommendations are specific and unambiguous; different options for management are clearly presented; key recommendations are easily identifiable; and the guideline is supported with tools for application.
5. APPLICABILITY: Potential organizational barriers in applying the recommendations are discussed; potential cost implications are considered; and the guideline presents key review criteria for monitoring and/or audit purposes.
6. EDITORIAL INDEPENDENCE: Externally funded guidelines should state explicitly that views and interests of the funding body have not influenced final recommendations; all group members should explicitly state potential conflicts of interest, which are recorded in the guideline.

The Conference on Guideline Standardization (COGS) checklist is another tool that specifies characteristics of a valid and usable clinical practice guideline.<sup>14</sup> In contrast to the AGREE instrument, which assesses guidelines after completion, the COGS checklist can be used during devel-



**Table 2**  
**Characteristics of a quality clinical practice guideline**

Topic	Description
1. Overview material	Structured abstract including release date, status, print and electronic sources
2. Focus	Primary disease/condition and intervention/service/technology
3. Goal	Goal guideline is expected to achieve, including rationale for topic
4. Users/setting	Intended users of the guideline and practice settings
5. Target population	Patient population eligible for guideline plus exclusion criteria
6. Developer	Organization(s) responsible for development plus author names/credentials
7. Funding source	Who sponsored development, what was role, what are conflicts of interest
8. Evidence collection	Literature search methods, including dates, databases, and filter criteria
9. Grading criteria	Method for grading recommendation strength and rating evidence quality
10. Evidence synthesis	How evidence was used to create recommendations
11. Prerelease review	How guideline developer reviewed and/or tested guidelines prior to release
12. Update plan	Expiration date for guideline and plans for updating
13. Definitions	Defines unfamiliar terms and those critical to correct application
14. Recommendations and rationale	Recommended actions are stated precisely with specific circumstances under which to perform them, and explicit linkage to supporting evidence
15. Benefits and harms	Potential benefits and risks associated with recommendations
16. Patient preferences	Role in decisions with substantial personal choice or values
17. Algorithm	Graphical description of the stages and decisions in clinical care
18. Implementation	Anticipated barriers to implementation, auxiliary materials, review criteria

Adapted from the Conference on Guideline Standardization.<sup>14</sup>

opment to improve quality. The 18 characteristics in the COGS checklist are shown in Table 2.

Guidelines meeting certain quality standards are included in the National Guideline Clearinghouse (NGC) database, an initiative of the Agency for Healthcare Research and Quality. NGC inclusion criteria are:<sup>15</sup>

1. The clinical practice guideline contains systematically developed statements that include recommendations, strategies, or information that assists physicians and/or other health-care practitioners and patients make decisions about appropriate health care for specific clinical circumstances.
2. The clinical practice guideline was produced under the auspices of medical specialty associations; relevant professional societies; public or private organizations; government agencies at the federal, state, or local level; or health-care organizations or plans. A clinical practice guideline developed and issued by an individual not officially sponsored or supported by one of the above types of organizations does not meet the inclusion criteria for NGC.
3. Corroborating documentation can be produced and verified that a systematic literature search and review of existing scientific evidence published in peer-reviewed journals was performed during the guideline development. A guideline is not excluded from NGC if corroborating documentation can be produced and verified detailing specific gaps in scientific evidence for some of the guideline's recommendations.
4. The full-text guideline is available upon request in print or electronic format (for free or for a fee), in the English language. The guideline is current and the most recent version produced. Documented evidence can be pro-

duced or verified that the guideline was developed, reviewed, or revised within the last five years.

Equally important to understanding what guidelines *are* is a clear appreciation of what guidelines *are not*. Without this perspective clinicians may become apprehensive about the impact of guidelines on their lives, and organizations may apply guidelines to situations for which they were never intended.

- Guidelines are *not* reimbursement policies.
- Guidelines are *not* performance measures.
- Guidelines are *not* legal precedents.
- Guidelines are *not* measures of certification or licensing.
- Guidelines are *not* for provider selection or public reporting.
- Guidelines are *not* recipes for cookbook medicine.

Guidelines are never intended to supersede professional judgment; rather, they may be viewed as a relative constraint on individual clinician discretion in a particular clinical circumstance.<sup>16</sup> Clinicians should always act and decide in a way that they believe will best serve their patients' interests and needs, regardless of guideline recommendations. Guidelines simply represent the best judgment of a team of experienced clinicians and methodologists addressing the scientific evidence for a particular topic.

Guidelines differ from *systematic reviews* and *evidence reports* that identify and combine studies using explicit methods to reduce bias, but do not typically define appropriate actions or incorporate values. In contrast, a guideline uses information from evidence reviews and

other sources to make specific recommendations by considering values and linking the strength of recommendation to the quality of evidence.

Last, evidence-based clinical practice guidelines are *not* intended for cost control or health-care rationing. Guidelines seek to produce optimal health outcomes for patients, minimize harm, and reduce inappropriate variations in clinical care. Whereas some of these outcomes may also reduce costs, financial benefits alone are generally not the main focus of an evidence-based clinical practice guideline.

## PRINCIPLES OF GUIDELINE DEVELOPMENT

Without substantial advance planning, guideline development is likely to be biased and inefficient. Moreover, an *a priori* protocol is mandatory to ensure attention to the COGS and AGREE quality standards. Based on literature review and direct experience in North America and the United Kingdom, Shekelle and colleagues<sup>17</sup> concluded that five steps are involved in the initial development of an evidence-based guideline:

1. identifying and refining the subject area
2. convening and running guideline development groups
3. assessing evidence identified by systematic literature review
4. translating evidence into recommendations
5. subjecting the guideline to external review

Turner and co-workers<sup>11</sup> compared approaches to guideline development in six handbooks from the Council of Europe, the World Health Organization, and national organizations in Australia, Scotland, New Zealand, and the United Kingdom. All handbooks agreed that key aspects of development included a multidisciplinary panel, consumer involvement, identifying clinical questions or problems, systemically reviewing and appraising the literature, a process for drafting recommendations, external consultation and review, and planned updating.

Guyatt and colleagues<sup>18</sup> have focused on grading evidence quality and recommendation strength in guidelines, emphasizing that both are separate and distinct processes essential to validity. An optimal grading system is characterized by simplicity and transparency for the clinician consumer, sufficient (but not too many) categories, explicitness of methodology for guideline developers, simplicity for guideline developers, consistency with general trends in grading systems, and an explicit approach to different levels of evidence for different outcomes.

The guideline development process described in this manual addresses the above issues, yet strives for a balance between rigor and pragmatism that maintains efficiency.

Efficiency is critical in guideline development, because moving from planning to completion in about 12 months helps avoid a situation in which new evidence continues to appear. With an efficient protocol in place, an organization can stagger guidelines under simultaneous development to result in a finished product every six months (depending on resources). The timeline in Table 1 has been developed to ensure rigor in development while promoting efficiency. The remainder of this manual describes the steps listed in Table 1 in terms of general concepts and specific suggestions based on prior experience.

## PLANNING

### Define Topic

Guidelines can be developed for a wide range of topics, including conditions (sinusitis, ear infections), procedures (tonsillectomy, tympanostomy tubes), and signs or symptoms (cough, hoarseness). Topics selected for guideline development should be high-priority and feasible.

*High-priority* topics have the potential for evidence-based practice to improve health outcomes, minimize undesirable variations in care, and reduce the burden of disease and health disparities. The Institute of Medicine has identified the following priority-setting criteria as common to most international guideline development groups:<sup>8</sup>

1. **DISEASE BURDEN.** Extent of disability, morbidity, or mortality imposed by a condition, including effects on patients, families, communities, and society overall.
2. **CONTROVERSY.** Controversy or uncertainty around the topic and supporting data.
3. **COST.** Economic cost associated with the condition, procedure, treatment, or technology related to the number of people needing care, unit cost of care, or indirect costs.
4. **NEW EVIDENCE.** New evidence with the potential to change conclusions from prior assessments.
5. **POTENTIAL IMPACT.** Potential to improve health outcomes and quality of life; improve decision making for patient or provider.
6. **PUBLIC OR PROVIDER INTEREST.** Consumers, patients, clinicians, payers, and others want an assessment to inform decision making.
7. **VARIATIONS IN CARE.** Potential to reduce unexplained variations in prevention, diagnosis, or treatment; the current use is outside the parameters of clinical evidence.

*Feasible* topics have a sufficient base of high-quality published evidence (ideally, randomized controlled trials) to drive guideline development, have one or more existing systematic reviews or meta-analyses already published on relevant issues, and have relatively clear definitions of the condition or procedure under consideration.

A steering committee that includes organizational leadership and broad stakeholder representation can help identify, prioritize, and refine guideline topics. Diversity of expertise and perspective helps minimize bias caused by conflicts of interest.

The AAO-HNS convened the Guideline Development Task Force as a steering committee for developing evidence-based guidelines and related knowledge products.<sup>19</sup> The task force includes representatives of all subspecialty groups within otolaryngology and of all relevant internal Academy groups, including research, patient safety, quality improvement, Board of Governors, and evidence-based medicine. Topics are solicited with a standardized form, based on principles outlined above, then presented to the task force for ranking and prioritization.

### Convene the Guideline Working Group

Perhaps the most important decision in creating a successful guideline relates to composition of the working group. A group size of 15 to 20 members encourages diversity and efficiency, yet is small enough to avoid delays and redundancy.

The group should consist of 1) the chair and two assistant chairs, 2) staff lead and assistant, 3) technical consultant, 4) content experts, 5) stakeholders from all relevant disciplines, including nursing, primary care, and allied health, and 6) a consumer representative. The roles and responsibilities of group members are outlined in the sections that follow.

### Identify Organizational Leadership

A *staff lead* is assigned as the primary liaison for the group, with one or more assistants who have the dual responsibility of supporting the lead and learning the process so they may serve as a future lead. Qualifications for staff lead include service as an assistant staff lead on a prior guideline panel, experience conducting literature searches and using a citation database, and a basic understanding of study design, medical terminology, and levels of evidence.

Specific responsibilities of the staff lead and assistants include:

1. conducting a preliminary search to assess topic feasibility
2. identifying guideline group members by working with internal leadership and relevant external organizations

3. scheduling and handling logistics for all conference calls and group meetings
4. working with the chair to create agendas and pre-distribute supporting materials
5. coordinating literature searches, organizing search results, and obtaining full-text articles
6. appraising the guideline for implementability using pre-determined methods
7. identifying external peer reviewers and collating comments for distribution to the chair
8. assisting the chair in developing and obtaining permissions for tables and figures
9. proofreading the guideline final draft, including checks for grammar and spelling
10. submitting a summary of key action statements and supporting evidence profiles for review and approval by the organizational board of directors
11. assisting the chair in formatting the final document for publication submission
12. obtaining copyright transfer and financial disclosure forms from working group members

A *technical consultant* is assigned to ensure that the working group adheres to methodologic standards and protocols endorsed by the organization, and to serve as a facilitator who supports the chair during conference calls and meetings. The technical consultant should be fluent with guideline methodology, understand the process of systematic review, and have direct experience with prior guidelines developed by the organization.

Developing valid guidelines is not intuitive, but is an acquired skill that is independent from clinical expertise and accomplishment. Whereas an explicit and comprehensive manual aids the process, it cannot substitute for hands-on experience.

### Identify Clinical Leadership

A *chair* should be identified to lead the group in developing the guideline and to work with the technical consultant and staff lead to ensure adherence to methodologic standards. The chair also facilitates the interpersonal aspects of the group processes, so the members work in a spirit of collaboration with balanced contribution from all members.

Specific responsibilities of the chair include:<sup>20,21</sup>

1. assisting the staff in planning conference calls and meetings
2. steering discussions according to the agenda
3. encouraging all members to contribute to discussions and activities of the group
4. remaining aware and constantly attentive to small-group processes, including how the group interacts, communicates, and makes decisions

5. establishing a climate of trust and mutual respect among members while remaining sensitive to preexisting inter-professional tensions and hierarchies
6. maintaining a unified group discussion free of subconversations and dominance
7. encouraging constructive debate without forcing agreement
8. winding up repetitive debate and disagreements through careful negotiation
9. summarizing main points and key decisions of a debate
10. delegating writing assignments and integrating completed assignments and group feedback into the draft guideline

The chair is appointed by a selection panel that includes organizational leadership, steering committee representation, the guideline staff lead, and the technical consultant. An ideal chair should be efficient and motivated, have demonstrated leadership ability, have prior experience with evidence-based guideline development, have demonstrated skills in scientific writing, and be fluent with using the internet, e-mail, and e-mail attachments. Candidates for chair will be asked to submit a curriculum vitae and declaration of competing interests, and to confirm that they understand and accept the substantial time commitment involved.

The chair should ideally *not* be a content expert for the guideline topic, but should be familiar with the scientific literature and management of the clinical condition. Content experts are usually abundant in an organization and can be readily added to the working group to fill in knowledge gaps. Conversely, the chair should be an impartial leader who stimulates discussion, not an advocate who injects his or her own opinions.<sup>17</sup>

One or two *assistant chairs* should be identified who will be asked to chair the next guideline development effort. To maintain a pipeline of guideline projects, a continuing source of leadership for upcoming projects is needed. The best way to groom new chairs is to have them serve on one or two prior guideline groups to learn methodology and expectations early on. An ideal assistant chair should have experience with evidence-based medicine, but does not necessarily need prior guideline development experience.

The chair is ultimately responsible for moving along the guideline process and keeping the group focused and task-oriented. Having more than one chair is inadvisable, because responsibilities can be easily shifted and diffused. Instead, the structure should include one chair and one or more assistant chairs, as noted above.

### Identify Partner Organizations

Guideline development panels should include individuals from a range of relevant stakeholder groups to minimize bias. Multidisciplinary participation helps identify and evaluate all relevant evidence, builds support among the intended guideline users, and increases the chances of addressing practical problems related to implementation.<sup>10</sup>

Many guidelines warrant input from nursing, consumers, and primary care clinicians. Based on the target population and setting, the working group may include internists, pediatricians, geriatricians, family practitioners, and emergency medicine physicians. Additional specialty clinicians are recruited as dictated by the specific topic or condition under study. Allied health professions are similarly recruited, and may include audiologists, physical therapists, speech-language pathologists, and others.

An excellent source of consumer participants for guideline development is Consumers United for Evidence-based Healthcare (CUE), a national coalition of health and consumer advocacy organizations, which empowers consumers through critical appraisal of articles, guidelines, and systematic reviews.<sup>22</sup> CUE is a project of the U.S. Cochrane Center and works closely with the Cochrane Consumer Network.

If another discipline is to be a *full partner* in developing the guideline, that discipline is approached early to secure interest and cooperation. Alternatively, working group members can be selected to represent their “discipline,” not their “organization.” In this model a pediatrician member of the working group would provide essential input for pediatrics as a discipline, but would not necessarily represent the American Academy of Pediatrics or imply their specific endorsement of the resulting guideline.

### Identify Guideline Working Group Members

In deciding what disciplines other than otolaryngology to include in guideline development, a useful approach is to ensure that *every discipline or organization that would be involved with implementation, including consumers, has a voice at the table*. This will nearly always include one or more primary care clinicians, since invariably they will be involved in counseling the patient and coordinating care with the specialist. Representatives of all relevant medical specialties other than otolaryngology must also be considered.

A single specialty group will reach different conclusions than a multidisciplinary group when presented with the same evidence.<sup>17</sup> Individuals from a single discipline are often biased towards procedures in which they have a vested interest. Involving multiple



disciplines tends to balance bias and produce more valid guidelines.<sup>8</sup>

Potential members of the working group can be identified by organizational leadership, partner organizations, the working group chair, and the staff liaisons. An understanding of evidence-based medicine is desirable. Individuals are invited as representatives of their field or discipline, but need not be content experts for the guideline topic. Content experts should be a minority voice on the working group to limit bias.

Specific responsibilities of the working group members include:

1. participating in all conference calls
2. attending all meetings with a commitment to teamwork and clear communication
3. reading all relevant materials and providing constructive comments and feedback during and between meetings
4. checking and responding to e-mails on a regular basis
5. completing personal assignments to meet deadlines
6. maintaining confidentiality
7. disclosing fully any potential conflicts of interest

The importance of choosing an appropriate working group cannot be overemphasized. This is called a “working” group for a reason: producing a guideline requires substantial time and effort. All members have a responsibility to other participants to behave with integrity, commitment, and a fully professional demeanor.

Despite the up-front commitment of all working group members to participate fully in guideline development, conflicts or unexpected circumstances may arise that threaten validity if an important discipline is not represented. Therefore, certain disciplines, which include primary care and selected others depending on the topic, should be represented by two group members to ensure representation.

### Compile Contact Information Grid

The staff lead should compile a grid of contact information for all working group members and organizational representatives. Included in the grid should be 1) name and degrees, 2) working group role, 3) organizational affiliation, 4) clinical and academic titles, 5) mailing address, 6) disclosed conflicts of interest, and 7) contact information.

### Conflict of Interest Disclosure

A conflict of interest exists when a participant or the participant’s institution has financial or personal relationships

with other people or organizations that may inappropriately influence (bias) his or her actions.

Despite good intentions, it is *not* appropriate for individuals to decide if a particular relationship causes conflict; their role is to declare, not interpret. The group as a whole must ultimately determine if a conflict may result in bias, and whether or not the degree of conflict excludes the individual from participating in the entire guideline or selected sections.

Financial relationships are easily identifiable, but conflicts can also occur because of personal relationships, academic competition, or intellectual passion. Examples of financial conflicts include employment, consultancies, stock ownership, honoraria, paid expert testimony, patents or patent applications, and travel grants. Full disclosure is advised regardless of whether the participant considers the relationship relevant to the guideline content.

The contact and disclosure list should be distributed to all members for verification and should be updated, as needed, during guideline development and prior to publication.

### Determine Dates for Conference Calls and Meetings

Adhering to a predetermined, specific timeline allows publication of the guideline within 18 months. Arranging dates for conference calls and meetings is particularly difficult when dealing with individuals representing multiple organizations and disciplines. Therefore, it is critical to plan early in the process. Events are planned using the timetable in [Table 1](#):

- Conference call #1 takes place in month 2.
- Conference call #2 takes place about four weeks later, in month 3.
- In-person meeting #1 takes place about four weeks later, in month 4.
- In-person meeting #2 takes place about six to eight weeks later, in month 6.

Conference calls are often most feasible if planned to start at 8:00 p.m. Eastern Standard Time. Calls should be generally scheduled for two hours. In-person meetings can begin at noon with a light working lunch to allow attendees to fly in the same morning. Similarly, they can end by noon the next day to allow a return flight the same day. A group dinner should be planned the first day. A convenient schedule is to begin on either Friday or Sunday, and end the next day.

The staff lead prepares a grid of potential dates for the calls and meetings. The grid is circulated by electronic mail to the chair, assistant chair, and technical consultant to determine available dates for the first two conference calls. For the in-person meetings and future conference calls, the grid may be circulated to the entire working group to assess availability. There will clearly be a need for compromise by some group members, since the odds of finding dates agree-

able to all are extremely low. Group members must commit to attending these meetings at the start.

The importance of having *all* working group members participate in *all* conference calls and attend *all* meetings cannot be overemphasized. Advance planning is the best guarantee of success, since maximal time is available for group members to adjust their schedules as needed and block out event dates in their calendars. If a group member cannot make this commitment, an alternate should be found as soon as possible.

## IDENTIFYING EVIDENCE

### Purpose

The validity of an evidence-based guideline depends in large part on an unbiased and comprehensive literature search. The goal is to locate the best evidence from all relevant sources, producing a comprehensive body of evidence that will allow clinical questions to be answered and highlight gaps in the evidence base where formal consensus methods may be needed.<sup>20</sup>

### The Role of Evidence in Guideline Development

Although identifying evidence is essential for guideline development, we suggest the proper role is as supporting cast, not protagonist:

- **EVIDENCE AS PROTAGONIST MODEL.** Many organizations publish “practice parameters” or “evidence-based reviews” as their primary quality products, having the literature search take center stage, with exhaustive evidence tables or textual discussions that rank and summarize citations. If recommendations are made, the strength is linked directly to level of evidence, sometimes with a threshold number of minimum studies of a specific level or combinations thereof, rather than an explicit consideration of benefits, risks, harms, and costs. Recommendations have uncertain validity because there is no systematic process to incorporate costs, harms, adverse events, uncertainty, vagueness, working group values, or patient preference. Moreover, recommendations become difficult to make when evidence gaps exist.
- **EVIDENCE AS SUPPORTING CAST MODEL.** An alternative approach, described in this manual, is to drive guideline development with considerations of quality improvement, using the literature search as one of many factors to translate evidence into action. In this model the ratio of benefits to harms and costs is considered equal to, or even greater than, level of evidence in formulating recommendations. Evidence profiles are used to state explicitly how values, patient preferences, and conflicts of interest were

incorporated. Recommendations are still possible with evidence gaps, but strength will be limited.

Although it is tempting to exclude topics with limited evidence from guideline development, it is precisely such topics that benefit most from inclusion because of uncertainty and conflicting opinions. Even if evidence is limited, recommendations are still possible if well-documented benefit or harm is identified.

Using expert opinion or consensus to fill evidence gaps is entirely appropriate, provided this basis is explicit and transparent to the critical reader.<sup>9</sup> Discussing topics with limited evidence allows guideline developers to highlight future research needs, with critical suggestions on how to best fill existing gaps. The guideline as a whole, however, should focus on topics with high-quality evidence and avoid over-reliance on expert opinion as a primary decision-making strategy.<sup>23</sup>

### Literature Search Stages

Similar to the National Institute of Clinical Excellence,<sup>20</sup> we have found that searching is an iterative process that is best implemented in three stages. The stages correspond to different phases of guideline development, and are discussed in detail at the appropriate point in the manual. The three stages of searching (Table 1) can be briefly summarized as:

1. identifying systematic reviews and clinical practice guidelines, performed before the first conference call
2. identifying randomized controlled trials, performed before the second conference call
3. identifying supplementary literature, performed after the first in-person meeting

All search stages must be documented for transparency and reproducibility. Specific considerations include databases, time periods, key words, subject headings, language restrictions, use of gray literature (eg, symposium proceedings), and selection criteria, such as filters, algorithms, or inclusion and exclusion criteria. A balance of pragmatism and rigor is required to avoid delays in the development process.

### Evidence Quality Assessment

Simply identifying reviews, guidelines, and randomized trials does not ensure quality, and basing decisions on research with weak design or flawed methodology may yield biased or invalid conclusions. Therefore, to filter out potentially biased or poorly conducted studies, quality assessment must be performed as part of identifying evidence. Suggestions

for assessing reviews, guidelines, and randomized trials are presented later in the manual when the related search stage is discussed.

## Performing Systematic Review and Meta-Analysis

An organization may find it necessary to perform a systematic review as part of guideline development if there are no published reviews, or if existing reviews are outdated or of poor quality. Systematic review is a rigorous and complex undertaking that often requires additional expertise, resources, and staff support.

All systematic reviews should be conducted using *a priori* protocols that adhere to standards for the conduct and reporting of meta-analyses, as suggested in the QUOROM statement for randomized trials<sup>24</sup> and the MOOSE statement for observational studies.<sup>25</sup> Systematic reviews can be used to define natural history using placebo group outcomes and the absolute or comparative efficacy of interventions.<sup>26,27</sup>

### Key Points to Remember

1. Identifying evidence involves a three-stage literature search, with different stages occurring at different times in the development process.
2. The literature search supports guideline development, not vice versa.
3. All aspects of searching must be documented for transparency and reproducibility.
4. Constant vigilance is required to balance rigor vs pragmatism so that guideline development is not stalled or delayed because of overly complex search strategies.

## STAGE 1 LITERATURE SEARCH

### Purpose

The stage 1 search establishes a foundation for the first working group conference call by identifying existing systematic reviews and practice guidelines related to the current topic. This provides important perspective on what has already been accomplished, what areas of controversy exist, how robust the evidence base is to support guideline development, and where the greatest opportunities lie for improving upon the existing knowledge base.

The stage 1 search is coordinated by the staff lead before the first working group conference call, using parameters defined by the chair and technical consultant. Search results are reviewed by the chair to eliminate irrelevant items. A summary grid is compiled and full-text files are obtained for distribution to the working group.

### Identifying Systematic Reviews

Systematic reviews will greatly facilitate guideline development because they identify and synthesize evidence in a

format that is readily usable by the working group. Systematic reviews and meta-analyses are found by:

- identifying Cochrane Reviews and Protocols via the Cochrane Collaboration website ([www.cochrane.org](http://www.cochrane.org)) and by searching the Cochrane Database of Abstracts of Reviews of Effects (DARE);
- locating evidence reports sponsored by the Agency for Healthcare Research and Quality (AHRQ) ([www.ahrq.gov](http://www.ahrq.gov)), which are often developed to support guideline development and may contain useful evidence tables, systematic reviews, and meta-analyses;
- searching standard databases—MEDLINE, EMBASE, and Cumulative Index to Nursing and Allied Health Literature (CINAHL)—using “systematic review” or “meta-analysis” as a publication type or text word in the title or abstract;
- using search filters of known validity for identifying systematic reviews;<sup>28</sup> and
- searching clinical evidence available from the BMJ publishing group ([www.clinicalevidence.bmj.com](http://www.clinicalevidence.bmj.com)).

### Identifying Clinical Practice Guidelines

Clinical practice guidelines may already exist for the topic under consideration, but do not preclude further guideline development. Existing guidelines may be outdated or may not have been developed with the methodologic rigor or relevancy that is currently sought. These documents, however, are a useful starting point for group discussions. Clinical practice guidelines can be identified by:

- searching the National Guidelines Clearinghouse ([www.guideline.gov](http://www.guideline.gov)), an initiative of AHRQ that serves as a public resource for evidence-based guidelines;
- searching the database maintained by the Guidelines International Network ([www.g-i-n.net](http://www.g-i-n.net)), which includes guidelines, evidence reports, and systematic reviews; and
- searching standard electronic databases for “guideline” or “practice parameter” as a text word in the title or abstract.

### Assessing Quality

*Systematic reviews* published by the Cochrane Collaboration or government agencies (AHRQ) are typically of high methodologic quality and may not require further assessment. Conversely, reviews authored by individuals or other organizations are highly variable in rigor and quality. Minimum quality criteria for systematic reviews might include 1) an *a priori*, hypothesis-driven protocol, 2) explicit and systematic literature search, 3) validated data extraction from source articles, 4) data pooling with standard statistical techniques, and 5) tabular presentation of results with graphical summaries.

*Clinical practice guidelines* are highly variable in quality regardless of origin. Minimum quality criteria might include 1) explicit scope and purpose, 2) multidisciplinary stakeholder involvement, 3) systematic literature review, 4) explicit system for ranking evidence, and 5) explicit system for linking evidence to recommendations.

## CONFERENCE CALL #1: DEFINING SCOPE

### Purpose

The first conference call (Table 1) sets the stage for guideline development by introducing working group members, defining the guideline timeline and scope, discussing conflicts of interest, and planning for the stage 2 literature search. The call is planned to last 120 minutes and may be recorded for future reference.

The staff lead records minutes of the call, for dissemination and review by the group after the call concludes. The main purpose is to document process, workflow, and decisions made, thereby avoiding the discussion of settled controversies. The group chair takes additional notes during the call to record ideas, concepts, definitions, and key phrases that may later prove difficult to reproduce or remember.

### Predistribute Electronic Materials

Documents should be distributed by e-mail prior to the conference call for review by participants before the call. Materials specific to the guideline that should be distributed include:

- agenda for the conference call
- working group contact information grid with conflict of interest disclosures
- summary grid of relevant systematic reviews and guidelines identified in the stage 1 literature search

General materials that should be distributed include:

- a copy of one or more recently published guidelines from the sponsoring organization to serve as a model of how the finished product will look. Suggested guidelines sponsored by the AAO-HNS include guidelines on otitis media with effusion,<sup>29</sup> acute otitis externa,<sup>3</sup> adult sinusitis,<sup>4</sup> cerumen impaction,<sup>5</sup> and benign paroxysmal positional vertigo.<sup>6</sup>
- reporting checklist from the Conference on Guideline Standardization (COGS)<sup>14</sup>
- article by Choudhry and co-workers<sup>30</sup> about conflict of interest disclosure

### Review Contact Information, Titles, Organizations

Group members should review contact information and titles for accuracy. Group members should briefly introduce themselves, including their areas of expertise and experience in developing prior guidelines, and their role in the workgroup. The need for any additional group members should be discussed, taking care to be sure that all relevant disciplines are adequately represented.

### Introduce Purpose, Methodology, Timeline

The purpose of sponsoring organization(s) in developing the guideline should be specified, and can be revised and updated as development proceeds. The purpose can often be divided into two distinct but related components:

1. objective, ie, general goals that implementation of the guideline are intended to bring about; and
2. rationale, ie, reasons for developing recommendations, including why the guideline is needed (eg, evidence of practice variation or inappropriate practice).

Here is an example of how purpose was stated in the AAO-HNS guideline on acute otitis externa: “The primary purpose of this guideline is to promote appropriate use of oral and topical antimicrobials for diffuse acute otitis externa and to highlight the need for adequate pain relief. Additional goals are to make possible an acute otitis externa performance measure and to make clinicians aware of modifying factors that can or may alter management (eg, diabetes, immune compromised state, prior radiotherapy, tympanostomy tube, non-intact tympanic membrane).”<sup>3</sup>

As another example, consider this statement of purpose from the AAO-HNS guideline on adult sinusitis: “The primary purpose of this guideline is to improve diagnostic accuracy for adult rhinosinusitis, reduce inappropriate antibiotic use, reduce inappropriate use of radiographic imaging, and promote appropriate use of ancillary tests that include nasal endoscopy, computed tomography, and testing for allergy and immune function.”<sup>4</sup>

After the guideline purpose has been discussed, the consultant provides a very brief overview of development methodology. Points worthy of emphasis include:

- The guideline will meet or exceed reporting standards defined by the COGS statement and AGREE instrument.<sup>13,14</sup>
- The guideline will be developed using an explicit, evidence-based process that incorporates group values and patient preferences.
- The process will involve three conference calls and two in-person meetings.
- The goal is to produce a document with key action statements for clinicians, highlighting the key areas of behavior change and quality improvement defined by the working group that address the stated purpose.
- The guideline will incorporate a series of about 10 to 18 key action statements, each of which is followed by amplifying text on why the statement is made and how the recommendation may be carried out, and a summary evidence profile.
- Each key action statement will have an associated strength of recommendation based on the quality and consistency of supporting evidence, plus a consideration of the benefit-harm relationship for any interventions.
- The document will conclude with implementation considerations and recommendations for future research.

The detailed methodology for classifying recommendations *should not* be discussed



at this time to avoid an unnecessary tangent. Instead, this is optimally discussed before classifications are made at the first or second in-person meeting. Members who want additional details can be referred to the American Academy of Pediatrics Policy Statement on Classifying Recommendations.<sup>16</sup>

### Discuss Conflicts of Interest

Group members must disclose all industry relationships and potential conflicts of interest during the first conference call. The group will then decide if any particular relationships are significant enough to preclude participation of any individual(s). Relationships should be thoroughly documented and included in the guideline manuscript.

Since more than 80 percent of guideline authors, in general, have potential conflicts of interest, the existence of a relationship alone is not sufficient to preclude participation.<sup>30</sup> Participants are only excluded if the nature of the relationship is considered by the group to interfere with objective participation (eg, equity relationship, patent holder, royalty arrangements). Based on the nature of the disclosed relationship, a member may be asked to not participate in a specific section of the guideline where a conflict may produce bias.

### Determine Guideline Scope

A well-crafted guideline has a clearly defined scope. Defining scope will occupy most of the first conference call, and may require a second for completion. Inexperienced guideline developers attempt to cover all aspects of a condition, resulting in a broad scope that will stall development efforts.

The key to progress is a razor-sharp focus from the start, recognizing that some issues important to some stakeholders will inevitably be left out.

*Define target condition or procedure.* The group should identify the conditions, procedures, or signs or symptoms for which the guideline is intended. This may be a single condition or a list of potential target conditions, which could be later condensed into those that can be realistically examined by the group within its allotted time. A guideline can be procedure-based instead of disease-oriented. For example, the emphasis can be on “tonsillectomy” as a procedure instead of tonsillitis as an acute or chronic condition.

Any diseases or procedures should be *explicitly* defined by the group. Definitions derived from publications on the topic can be used, if available, but a multidisciplinary group can often improve upon definitions advanced by an individual or single discipline. This is a particularly valuable contribution when existing definitions are controversial or unclear.

The definition of the target or procedure should be clear and concise (Table 3). The definition, however, should be distinguished from diagnostic criteria, which are typically specified later in the guideline and have more precise and detailed information to guide clinicians.

*Define target patient or clinical presentations.* The authoring group should specify the type of patient for whom the guideline is intended as precisely as possible. The target patient can be specified in terms of demographics, presenting signs and symptoms, past health history, results of previous diagnostic tests, or similar criteria.

Equally as important as defining the target patient is defining clearly the types of patients or clinical presenta-

**Table 3**  
Sample definitions from clinical practice guidelines

Guideline	Definition
Otitis media with effusion guideline <sup>23</sup>	Otitis media with effusion as discussed in this guideline is defined as the presence of fluid in the middle ear without signs or symptoms of acute ear infection. Otitis media with effusion is considered distinct from acute otitis media, which is defined as a history of acute onset of signs and symptoms, the presence of middle-ear effusion, and signs and symptoms of middle-ear inflammation.
Adult sinusitis guideline <sup>4</sup>	Rhinosinusitis is defined as symptomatic inflammation of the paranasal sinuses and nasal cavity. The term rhinosinusitis is preferred because sinusitis is almost always accompanied by inflammation of the contiguous nasal mucosa. Rhinosinusitis may be further classified by duration as acute (less than 4 weeks), subacute (4-12 weeks), or chronic (more than 12 weeks with or without acute exacerbations).
Cerumen impaction guideline <sup>5</sup>	Cerumen impaction is defined as an accumulation of cerumen that causes symptoms, prevents a needed assessment of the ear canal/tympanic membrane or audiovestibular system, or both. Although “impaction” usually implies that cerumen is lodged, wedged, or firmly packed in the ear canal, our definition of cerumen impaction does not require a complete obstruction.
Benign paroxysmal positional vertigo guideline <sup>6</sup>	Positional vertigo is defined as a spinning sensation produced by changes in head position relative to gravity. Benign paroxysmal positional vertigo is defined as a disorder of the inner ear characterized by repeated episodes of positional vertigo.

tions that are beyond the scope of the group's analysis. One or more exclusion criteria should generally accompany the definition. For example, consider this definition from the AAO-HNS guideline on acute otitis externa: "The target patient is aged 2 years or older with diffuse acute otitis externa, defined as generalized inflammation of the external ear canal, with or without involvement of the pinna or tympanic membrane. This guideline does not apply to children under age 2 years or to patients of any age with chronic or malignant (progressive necrotizing) otitis externa. Acute otitis externa is uncommon before age 2 years, and very limited evidence exists regarding treatment or outcomes in this age group. Although the differential diagnosis of the 'draining ear' will be discussed, recommendations for management will be limited to diffuse acute otitis externa, which is almost exclusively a bacterial infection. The following conditions will be briefly discussed but not considered in detail: furunculosis (localized acute otitis externa), otomycosis, herpes zoster oticus (Ramsay Hunt syndrome), and contact dermatitis."<sup>3</sup>

Here is another example of target patient definition from the AAO-HNS guideline on cerumen impaction: "The target patient for this guideline is over 6 months of age with a clinical diagnosis of cerumen impaction. The guideline does not apply to patients with cerumen impaction associated with the following conditions: dermatologic diseases of the ear canal; recurrent otitis externa; keratosis obturans; prior radiation therapy affecting the ear; previous tympanoplasty/myringoplasty or canal wall down mastoidectomy. However, the guideline will discuss the relevance of these conditions in cerumen management. The following modifying factors are not the primary focus of the guideline, but will be discussed relative to their impact on management: non-intact tympanic membrane (perforation or tympanostomy tube); ear canal stenosis; exostoses; diabetes mellitus; immunocompromised state; or anticoagulant therapy."<sup>5</sup>

*Define the intended audience and practice settings.* The decision about the intended users of the guideline needs to be made early in the process, since it influences decisions about the interventions that will be considered and the audiences to which the language in the final product and specific implementation suggestions will be directed. Ideally, a representative of each target audience group or organization should be included on the guideline working group. Stakeholder representatives should also be involved in reviewing and pre-testing the document.

Practice settings should also be defined, since a guideline may be applicable only in selected settings (eg, rural, primary care, hospital emergency room, operating room, managed care, specific geographic regions). The working group should identify those settings in which using the guideline would be appropriate as well as settings where it should not be applied.

Here is an example of how practice setting was defined in the AAO-HNS otitis media with effusion guideline: "The

guideline is intended for use by providers of health care to children, including primary care and specialist physicians, nurses and nurse practitioners, physician assistants, audiologists, speech-language pathologists, and child development specialists. The guideline is applicable to any setting in which children with otitis media with effusion would be identified, monitored, or managed."<sup>29</sup>

As another example, consider the definition used in the guideline on benign paroxysmal positional vertigo: "The guideline is intended for all clinicians who are likely to diagnose and manage patients with benign paroxysmal positional vertigo and applies to any setting in which benign paroxysmal positional vertigo would be identified, monitored, or managed."<sup>6</sup>

*Identify interventions to consider and exclude.* The group should generate a list of the clinical interventions (diagnostic tests, treatments, preventive measures) that will be considered in developing the guideline. This list should include all interventions relevant to the topic. A sample list developed for use in a sinusitis guideline is shown in Table 4.

The purpose of the topic list is to document transparency and to stimulate discussion as development proceeds, reminding the group of all interventions available for consideration. In contrast, the list is not intended as an outline or template for writing the guideline, since many items will be outside the document focus.

A similar list of exclusions should be generated. For example, some groups may be reluctant to evaluate drugs, procedures, or other interventions that have only recently been introduced into practice and have limited experience regarding long-term benefits and harms. Other groups may find these relevant. Any exclusions should be specifically noted in the list of interventions considered (Table 4).

*Identify outcomes to consider.* Outcomes should be selected prospectively that limit scope and provide measures against which to evaluate the effectiveness of the recommendation's limit.

- **HEALTH STATUS OUTCOMES** are direct measures of physical morbidity, emotional well-being, mortality, or some other health-related construct. Examples include audiometric hearing levels or survival rates for head and neck cancer (eg, two-year, five-year, and 10-year).
- **FUNCTIONAL HEALTH STATUS** measures reflect how a person functions physically, emotionally, and socially, with or without aid from the health-care system. There are many general and disease-specific surveys available to assess this construct.

**Table 4**  
**Sample list of interventions considered in guideline development for sinusitis**

Diagnosis	Treatment	Prevention
targeted history	watchful waiting/observation	topical steroids
physical examination	education/information	immunotherapy
anterior rhinoscopy	systemic antibiotics	nasal lavage
transillumination	topical antibiotics	smoking cessation
nasal endoscopy	oral/topical steroids	hygiene
nasal swabs	systemic/topical decongestants	education
antral puncture	antihistamines	pneumococcal vaccination
culture of nasal cavity, middle meatus, or other site	mucolytics	influenza vaccination
imaging procedures	leukotriene modifiers	environmental controls
blood tests: CBC, others	nasal saline	
allergy evaluation and testing	analgesics	
immune function testing	complementary and alternative medicine	
gastroesophageal reflux	postural drainage/heat	
pulmonary function tests	biopsy (excluded from guideline)	
mucociliary dysfunction tests	sinus surgery (excluded from guideline)	

Adapted from Rosenfeld et al.<sup>4</sup>

- **QUALITY-OF-LIFE** measures reflect how a person perceives his or her functional health status. Like patient satisfaction, this is an inherently subjective construct that can be measured with surveys.

Other measures to consider include cost, quality, and utilization. Often the outcome of interest is related only tenuously to the proposed interventions. In such cases, proxy indicators of outcome or process may be selected.

Here is an example of outcome definition from the AAO-HNS acute otitis externa guideline: “The primary outcome considered in this guideline is clinical resolution of acute otitis externa. Additional outcomes considered include minimizing the use of ineffective treatments; eradicating pathogens; minimizing recurrence, cost, complications and adverse events; maximizing the health-related quality of life of individuals afflicted with acute otitis externa; increasing patient satisfaction; and permitting the continued use of necessary hearing aids.”<sup>3</sup>

As another example, consider the definition from the cerumen impaction guideline: “The primary outcome considered in this guideline is resolution or change in the signs and symptoms associated with cerumen impaction. Secondary outcomes include complications or adverse events. Cost, adherence to therapy, quality of life, return to work or activity, return physician visits, and effect on comorbid conditions (eg, sensorineural hearing loss, conductive hearing loss) were also considered.”<sup>5</sup>

## Define Parameters for the Stage 2 Literature Search

The stage 2 literature search, described in the next section, identifies randomized controlled trials. During the first conference call, the parameters of the literature search conducted by the staff are discussed and defined.

The group should define any constraints on the initial literature search: published vs unpublished data, language restrictions (eg, English language only), time periods (eg, 1980 or later), or age groups (eg, adults, children, or both).

The group should discuss keywords for the search. If systematic reviews are already published (or in process at Cochrane), the search strategies used are reviewed for relevance to the current project. Suggestions for MeSH (medical subject heading) terms or other search keywords are solicited.

The group should discuss search strategies for randomized controlled trials that will be combined with the keywords to identify evidence.

The group should discuss bibliographic sources that will be used for the search, including the role of gray literature (eg, symposium proceedings).

## Working Group Assignments and Deadlines

At the end of the call the group reviews specific assignments or requests for additional information made during the call. Deadlines are assigned for completing the assignment, emphasizing the importance of responding within the time frame specified.

After the call the staff lead forwards notes and minutes to the chair for review and clarification. The revised minutes are distributed to the group for review and feedback. The definitions, scope, and purpose are further refined by e-mail exchange before the next conference call.

## Key Points to Remember

1. The stage 1 literature search for systematic reviews and clinical practice guidelines must be performed, assessed by the chair, and compiled into a summary grid before the call.

2. Most of the call will be spent defining the target condition or procedure, using clear, concise language that can be readily understood by all readers.
3. A sharply defined purpose and scope for the guideline will facilitate future efforts.
4. Emphasis is placed on broad concepts and group consensus; details and specifics are filled in after the call via e-mail exchange.

5. restricts the final set (1 or 2 or 3 or 4) by excluding “animals” as MeSH term.

## STAGE 2 LITERATURE SEARCH

### Purpose

The second step in identifying evidence is to assess the quantity and scope of randomized controlled trials (RCTs) available to support guideline development. Recommendations are strongest when supported by RCTs or systematic reviews of RCTs, and a paucity—or surplus—of quality studies may impact group decisions.

The stage 2 search should be coordinated by the staff lead before the *second* working group conference call, using parameters defined by the working group during the first conference call. Search results are reviewed by the chair and assistant chairs to eliminate irrelevant items. Remaining RCTs are organized by broad subject headings to facilitate group discussion and reference. A summary grid is compiled and distributed to the working group. The grid is most useful if some brief, descriptive information is included for each trial, such as sample size, blinding (open, single, or double), and industry funding (no or yes).

### Identifying Randomized Trials

Randomized trials are most valuable for evaluating therapeutic interventions. Different search strategies are required for questions related to prognosis, natural history, diagnostic tests, etc. Relevant clinical trials can be identified by:

- searching standard databases (MEDLINE, EMBASE, and CINAHL) for randomized controlled trial as a specific publication type;
- searching standard databases for “double blinds” or “randoms” in all fields;
- searching the Cochrane Central Register of Controlled Trials, accessible via the Cochrane Library website;
- manual cross-check of bibliographies from systematic reviews; and
- using other strategies or search filters of known validity, such as the Cochrane highly sensitive strategy for MEDLINE randomized trials that searches for:<sup>31</sup>
  1. “randomized controlled trial” or “controlled clinical trial” as publication type, or
  2. “randomized,” “placebo,” or “randomly” in the abstract, or
  3. “clinical trials as topic” as a Medical Subject Heading (MeSH) term, or
  4. “trial” in the title, and

### Assessing Quality

*Randomized controlled trials* are highly variable in methodology and validity. A simple and efficient scale can be used to rank quality from 1 (poor) to 5 (excellent) based on 1) method and adequacy of randomization, 2) method and adequacy of masking, and 3) reporting of withdrawals and dropouts.<sup>32</sup> A similar quality scale is available for randomized trials included in systematic reviews.<sup>33</sup>

## CONFERENCE CALL #2: IDENTIFYING TOPICS

### Purpose

The primary purpose of the second conference call (Table 1) is to refine and polish the concepts developed in the first call, particularly the scope and definition(s). The interval between the first and second conference calls should be kept short, about four to six weeks, to facilitate recall and sustain momentum.

The stage 2 literature search is now available and will help identify errors, omissions, and exclusions in the earlier discussion. The call ends with a discussion of quality improvement opportunities that are used to form a preliminary topic list, which will be further refined and prioritized by electronic mail exchange after the call. The call is planned to last 120 minutes and may be recorded.

### Predistribute Electronic Materials

Documents should be distributed by e-mail prior to the conference call for review by participants before the call. Materials for predistribution include:

- agenda for the conference call
- updated working group contact information grid with conflict of interest disclosures
- results of the stage 2 literature search for randomized controlled trials
- minutes of the first call, updated to include post-call assignments and e-mail exchanges
- any other information on content or methodology identified by working group members that would be helpful in educating the group or stimulating discussion

### Review First Conference Call and Timeline

Minutes from the first conference call are reviewed, with emphasis on the guideline purpose, scope, and definitions. Feedback is also solicited on the stage 2 literature search, especially from the group content experts, regarding content, organization, and possible omissions.

- Decisions made at the first meeting are reviewed and updated. Revisions to the definition(s), scope, or purpose are discussed in detail until consensus is achieved. Some



changes will likely occur, now that the group is more familiar with the topic and available evidence.

- It is essential to arrive at an unambiguous definition of scope. If issues arise that require additional discussion, they can be dealt with by e-mail after the call or, if needed, an additional conference call with all group members or a subgroup.
- The stage 2 literature search is briefly reviewed and discussed, beginning with a discussion of methodology by the staff lead and chair. Feedback from the content experts is important to ensure the search is comprehensive, well organized, and free of obvious omissions.
- The timeline is discussed, emphasizing dates for the in-person meetings, which should have been decided well in advance of this call.

### Begin Discussion of Topics for Key Action Statements

The heart of a guideline is a series of *key action statements* that reflect issues deemed most important by the group. Although there is no rigid guide as to the number of key statements, they are usually limited to about 10 to 18 based on the guideline scope. Since each statement will require supporting text and an evidence profile, the number is limited for feasibility and timeliness. The goal is to achieve maximal quality improvement with a manageable set of actions.

Quality health care ideally is patient-centered yet also accounts for the needs of the population. A useful definition of quality for guideline development is how well physicians and health-care institutions fulfill their care obligations to individual patients, and how well patients, physicians, and health-care institutions enable these obligations to be fulfilled justly across the population. The goal is to improve desired health outcomes that are consistent with current professional knowledge.<sup>34</sup>

The process for developing key action statements begins by asking the working group to suggest *topics* that represent opportunities for quality improvement within the guideline scope. A given topic may become the basis for a key action statement or, if deemed of lesser importance, may be incorporated into the supporting text of a related statement. Opportunities for quality improvement may be broadly summarized as:<sup>8</sup>

1. promoting appropriate care
2. reducing inappropriate or harmful care
3. reducing variations in delivery of care
4. improving access to care
5. improving the knowledge base across disciplines
6. educating and empowering clinicians and patients
7. facilitating coordination and continuity of care
8. facilitating ethical care

After reviewing the above list with the group, the chair solicits feedback for quality improvement topics. The points below should help in moving the discussion along:

- Begin by asking the group, “If we could only discuss a few aspects of this condition or procedure, what topics would we focus on to most improve quality of care and patient outcomes?” A related question is “What should we focus on to minimize harm?”
- Additional topics may be gleaned from the literature searches. Examining the scope of guidelines and systematic reviews from the stage 1 search is often very useful. Similarly, the availability of randomized controlled trials, identified in the stage 2 search, is a superb source of feasible topics to consider.
- Last, the group should review the intervention list from the previous conference call (Table 4) to ensure that no important aspects of management were omitted from the discussion.

Recall that the quality-driven approach allows all important topics to be included, even if the supporting evidence is weak or limited. Although recommendations are facilitated by strong evidence, important topics with weak evidence may still become key action statements if there is a clear preponderance of benefit or harm. The group should focus on potential quality impact in selecting topics, not primarily on level of evidence.

Topic suggestions should be short and simple, emphasizing content, not structure. No attempt should be made at this time to create polished action statements. The topic list is often much longer than the eventual list of key action statements. All group members must contribute to ensure multidisciplinary involvement. At least 20 to 30 topics are desired, roughly twice the number of anticipated key action statements.

### Compose Preliminary Topic List

As topics are suggested by working group members, the chair composes a simple list, assisted by the staff lead. The goal is to have a starting point for electronic exchange after the call that will refine and complete the entries.

A sample topic list developed from conference call #2 for a sinusitis guideline is shown below.

1. Initial management of acute sinusitis: observation vs antimicrobials
2. Optimizing diagnostic accuracy for acute sinusitis
3. What is an appropriate medical evaluation?
4. Antimicrobial selection for acute sinusitis: when to change, duration
5. Indications for imaging studies (plain radiographs, computed tomography) in sinusitis

6. Symptomatic relief for acute sinusitis, especially if antimicrobials are withheld
  7. Effectiveness and benefit of adjunctive measures for sinusitis
  8. Indications for allergy/immunology assessment
  9. Indications for nasal endoscopy in sinusitis
  10. Categorizing acute vs chronic vs recurrent acute: terminology issues
  11. Environmental control measures: smoking, environmental tobacco smoke, allergens
  12. Prevalence of antimicrobial resistance/resistant organisms for acute sinusitis
  13. Control of allergies as adjunctive therapy for sinusitis
  14. Preventing recurrent sinusitis
  15. Interpretation of imaging findings
  16. Patient education, information therapy
  17. Importance of identifying modifying factors or underlying conditions (eg, cystic fibrosis, prior surgery, immune deficiency, immotile cilia)
  18. Can sinusitis be managed over the phone (without office assessment)?
3. **DRAFTING THE HEALTH-CARE BURDEN SECTION.** The assistant chairs are asked to compose a brief summary of *health-care burden* for the guideline topic, about three to six paragraphs in length, which addresses issues of incidence, prevalence, direct cost, and indirect cost. This section will eventually be incorporated into the introductory text, and is a basis for stimulating interest by the media, public, and other stakeholders after publication.
  4. **SPECIFIC INFORMATION REQUESTS.** Any requests for additional information to clarify issues discussed during the call are reviewed and assigned to specific group members for action. The need for any supplemental literature searches is also discussed.

Deadlines are assigned for completing any assignments, emphasizing the importance of responding within the time frame specified. Group members are reminded of the dates for the upcoming in-person meetings.

### Key Points to Remember

1. The stage 2 literature search for randomized controlled trials must be performed, assessed by the chair and assistant chairs, and compiled into a summary grid before the call.
2. Much of the call will be spent refining and polishing the guideline scope and definitions.
3. A topic list is generated based on literature searches and on group input regarding opportunities for quality improvement.
4. Emphasis is placed on creating a preliminary topic list, which will be refined after the call and serve as a starting point for creating key action statements at the next meeting.

Many of the topic suggestions for sinusitis were the basis for key action statements, but the final wording had little relationship to how the topic first appeared. Topics are best viewed as the raw material for deriving key action statements, which is the subject of the first in-person meeting. The purpose of conference call #2, and the subsequent electronic exchange, is to create a robust platform of raw material to assist the group at the meeting.

### Working Group Assignments and Deadlines

At the end of the call the group reviews specific assignments or requests for additional information made during the call. Four types of assignments will follow the call:

1. **FINALIZING THE TOPIC LIST.** The preliminary topic list is sent to all working group members for revision and comment soon after the call. Members are requested to add any new topics, thought of after the call, to the existing list. A final topic list is compiled by the staff lead, which will be used to rank and prioritize topics as described in the next section.
2. **DRAFTING THE INTRODUCTION AND GUIDELINE PURPOSE.** The chair or his/her designee(s) should create a rough draft of the first two sections of the guideline, the Introduction and the Guideline Purpose, using prior guidelines as examples. The Introduction states why the topic was chosen, why it is important, precisely how the topic is defined (using language developed by the group), and to whom the guideline does—and does not—apply (scope). The section on Guideline Purpose states what the group seeks to achieve, what is the anticipated impact of the guideline on clinical care, and for what specific audience and situations the guideline is intended.

## PLANNING FOR IN-PERSON MEETING #1

### Logistics

The room should be large enough to comfortably accommodate the working group and facilitate discussion. One suggestion is to use a rectangular or U-shaped seating arrangement, with the chair and assistant chairs at the head, and participants along the sides.

A digital projector and screen are used for presentations (see below) and for real-time projection of meeting notes taken by the chair. The screen should be large enough to be readily seen in all parts of the room. Internet access should be available for literature searches and to address questions that arise during the meeting.

### Presentations

The first in-person meeting will ideally have several focused PowerPoint presentations to set the stage for the ensuing discussions. The following topics are suggested, allowing 30 minutes for each:

1. Overview of guideline methodology, presented by the consultant or methodologist, to familiarize members with the development process and show examples from other guidelines of what the group is seeking to accomplish
2. Current controversies in managing XYZ topic, presented by a content expert, to illustrate issues the group will be discussing, and to orient group members who may have limited knowledge of the guideline topic
3. Systematic reviews and guidelines about XYZ topic, presented by the staff lead, to summarize the stage 1 and stage 2 literature search findings

### Creating a Guideline Template

Progress at the first meeting is facilitated if the chair or staff lead creates a guideline template that will be updated in real time during the meeting. The template is based on the format used in previously published guidelines by the organization. An outline of the major sections in AAO-HNS guidelines is shown in [Table 5](#).

### Ranking Topics and Assigning Evidence

The final topic list, based on electronic exchange after the second conference call, is made into a two-column table. The first column, left blank, has the heading “Rank,” and the second column, containing the topics, has the heading

“Topic.” The order of the topics is not important, and can simply correspond to the sequence in which they were suggested by the group.

The staff lead distributes the topic list to working group members with the following instructions, replacing the number “31” in this example with the total number of topics: “Please rank the 31 topics below in order of importance for inclusion in this guideline by placing a number from 1 to 31 under the ‘Rank’ column. Please use each number only once. Assign the number ‘1’ to the most important topic, ‘2’ to the next most important, ‘3’ to the 3rd most important topic, so on and so forth. Number ‘31’ will be the least important topic. The table should *not be sorted*, but should be left in the original order. In addition, please send any comments to [staff lead e-mail address]. Thank you for your time.”

The rank lists are collated by the staff lead to determine the mean rank score for each topic, with lower scores indicating higher priority. A table is created with the items sorted by rank score with additional columns for the topic, number of systematic reviews relevant to the topic (based on the stage 1 search), and number of randomized trials relevant to the topic (based on the stage 2 search). An example of a completed topic rank list is shown in [Table 6](#). The literature search in compiling this table does not need to be exhaustive at this stage; it is simply intended as a guide to the evidence landscape for the upcoming meeting.

**Table 5**  
Suggested template for major sections in a clinical practice guideline

Section	Description
Abstract	Structured summary of objective, methods, purpose, and results
Introduction	Brief overview of key issues to engage the reader, including disease burden, target patient, definitions of key terms, and other introductory information
Guideline purpose	Explicit statement of why the group undertook the guideline, intended target audience, guideline exclusions, modifying factors (if applicable), and what the current effort adds beyond already published guidelines or literature
Health-care burden	Detailed information from primary sources on incidence, point prevalence, cumulative or lifetime prevalence, patient-based outcomes (eg, quality of life), direct and indirect costs, and outcomes considered in the guideline
Methods	Description of general methods, literature searches, and methodology for assessing levels of evidence, determining strength of recommendation, and handling issues related to competing interests or financial conflicts
Guideline key action statements	General description of the format with a summary table of all statements in the guideline, followed by individual listing of key action statements with corresponding supporting text and evidence profiles
Implementation consideration	Discussion of how the guideline will be disseminated, what anticipated implementation barriers will be encountered and how they will be handled, and what supporting materials will be developed for implementation
Research needs	Suggestions for future research to fill evidence gaps identified in the process of guideline development
Disclaimer	Legal disclaimer approved by the organization regarding guidelines
Acknowledgements	Recognition of individuals and their contribution to the guideline
Author information	Author affiliations and contributions to the manuscript
Financial disclosure	Full disclosure of competing interests for all authors, including those not deemed relevant to guideline content
References	Literature citations

**Table 6**  
**Topic list rankings from the AAO-HNS guideline on hoarseness**

Mean rank	Topic	Randomized trials	Systematic reviews
2.2	Role of history in diagnosis		
2.3	Role of physical examination in diagnosis		
6.8	How long to observe/treat empirically before laryngoscopy	2	
9.2	Method of examining the larynx		
8.2	Distinguishing between primary and secondary hoarseness		
9.7	Appropriate use of reflux medications for hoarseness	6	1
9.1	Evaluation for comorbidity: dysphagia, airway problems		
12.2	Need for radiography or imaging studies (CT, MRI)		
13.6	Duration of antireflux therapy and adverse effects		
14.0	Appropriate use of steroids in managing hoarseness		
14.7	Diagnostic testing: laryngoscopy in children		
16.1	Indications/role of voice therapy	5	4
16.0	Concurrent treatments: medications, therapy, surgery		
16.1	Appropriate use of antibiotics in managing hoarseness		
16.6	Indications for surgical intervention	3	1
16.8	Preventing hoarseness: hydration, vocal hygiene, smoking cessation, environmental changes	1	1
16.9	Role of voice rest (when/how long)		
17.3	Interdisciplinary approach to the hoarse patient: collaboration		
17.6	Professional voice users vs casual voice users	2	
19.6	Care of the geriatric voice; when to refer		
19.5	Role/timing of electromyography in diagnosis		
19.9	Interventions for children: observation, voice therapy, others		
21.0	Importance of visualizing the larynx in patients with hoarseness undergoing asthma therapy (inhaled steroids)	4	
20.2	Follow-up care: when to assess, how often; role of patient self-assessment and quality-of-life measures, instruments		3
20.7	Laryngeal papilloma as a cause of hoarseness in children		
23.9	Objective voice measures: acoustics and aerodynamics		
23.5	Cognitive impairment as a modifying factor		
22.5	Congenital anomalies as a modifying factor		
25.3	Recognize phono-traumatic hoarseness in children		
23.1	Botulinum toxin as a treatment for hoarseness	3	9
25.6	Role of complementary/alternative medicine		1

## UNDERSTANDING KEY ACTION STATEMENTS

A guideline can only be implemented if the recommendations are clear and identifiable. This goal is best achieved by structuring the guideline around a series of *key action statements*, which are supported by amplifying text, evidence profiles, and recommendation grade. Unfortunately, recommendation statements are often not readily identifiable in published guidelines and many statements are not executable as written.<sup>35</sup> This section introduces the concept of key action statements, laying the foundation for developing the statements from the topic list at the first in-person meeting.

### Crafting Valid Recommendation Statements

The group must understand the purpose and structure of key action statements before they can begin creating

them from the prioritized topic list. Key statements are action-oriented prescriptions of specific behavior from a clinician. As such, they should suggest *measurable* activities that can form the basis of performance measures or other quality initiatives.

Key statements are quality-driven and propose actions that will improve quality of care. These actions include, but are not limited to, reducing variations in care, improving diagnosis/recognition, promoting appropriate care, avoiding unnecessary tests or interventions, improved coordination of care, and improved patient safety.

An ideal key action statement describes:

- *when* (ie, under what specific conditions)
- *who* (specifically)
- *must, should, or may* (ie, the level of obligation)



- do *what* (precisely what actions)
- to *whom*

Key action statements should be brief, yet precise. The accompanying text amplifies *why* the recommendation is important and *how* it is to be carried out.

### Examples of Key Action Statements

Examples of key action statements from prior guidelines are listed below with the specific action requested in italics. These examples are *polished* statements developed after extensive discussion, and not necessarily what was initially proposed based on the topic list.

- Clinicians *should diagnose* cerumen impaction when an accumulation of cerumen is associated with symptoms, or prevents needed assessment of the ear, or both.
- Clinicians *should treat* the patient with cerumen impaction with an appropriate intervention, which may include one or more of the following: cerumenolytic agents, irrigation, or manual removal other than irrigation.
- Clinicians *should question* patients with benign paroxysmal positional vertigo for factors that modify management, including impaired mobility or balance, central nervous system disorders, a lack of home support, and increased risk for falling.
- Clinicians *should counsel* patients regarding the impact of benign paroxysmal positional vertigo on their safety, the potential for disease recurrence, and the importance of follow-up.
- Clinicians *should not obtain* radiographic imaging for patients who meet diagnostic criteria for acute rhinosinusitis, unless a complication or alternative diagnosis is suspected.

All of the statements above apply to “clinicians” as the “who,” followed by the word “should” then an action statement. The word “should” qualifies the strength of the statement (the subject of in-person meeting #2) and is replaced by “may” if the level of evidence is not strong or the harm-benefit relationship is unclear. Recommendations supported by consistent, high-quality evidence and a strong preponderance of benefit over risk, harm, and cost may occasionally be associated with the term “must.” Developers should understand the possible legal and reimbursement ramifications when using that term.

The final wording of key action statements will be determined based on evidence profiles at the next in-person meeting, so the group should not waste time debating whether “must,” “should,” or “may” is appropriate

at this point. At this stage the emphasis is on concepts and clarity, not precise wording.

## IN-PERSON MEETING #1: DRAFTING KEY ACTION STATEMENTS

### Purpose and Organization

The goal of this meeting to develop a “straw man” draft of the guideline’s key action statements based on the prioritized topic list. A consensus is reached regarding key statements, the messages to be delivered, and the order in which they are to be presented. The template resulting from this meeting facilitates working group assignments to sketch in the supporting text and other details prior to the next meeting. The task of classifying the statements into recommendations can be briefly discussed, but the process is best deferred until the next meeting.

As one of only two in-person meetings during guideline development, the venue must be organized efficiently. One suggestion is to begin at noon and end at 1:00 p.m. the next day, with two working lunches and a group dinner. The noon start allows some members to arrive the same morning, and the early afternoon finish allows all members to return the next day. In between, there is about 10 hours of working time, excluding the dinner. This time allotment, combined with effective leadership from the chair and consultant, should be adequate for completing the meeting agenda.

### Predistribute Electronic Materials

Documents should be distributed by e-mail prior to the conference call for review by participants before the meeting. Materials for predistribution include:

- agenda for the conference call
- updated working group contact information grid with conflict of interest disclosures
- minutes of the second conference call
- prioritized topic list with the number of relevant randomized trials or systematic reviews
- draft introductory sections of the guideline: introduction, purpose, and health-care burden
- selected full-text documents, including guidelines and systematic reviews, if available
- any other information on content or methodology identified by working group members that would be helpful in educating the group or stimulating discussion

### Presentations by Methodology and Content Experts

The prearranged PowerPoint presentations on methodology and content are given at the start of the meeting. Presentations should be informative, not didactic, focusing on ideas and concepts to stimulate the group, rather than attempting to make

definitive statements. Current methodology and common pitfalls in guideline development should be described.

### Reviewing the Draft Introduction, Purpose, and Health-care Burden

The introductory sections of the guideline, composed by the chair and assistant chairs after the second conference call, are reviewed and discussed by the group for broad concepts, clarity, and consistency. The first two sections are based largely on discussions at the first two conference calls and the content should accurately reflect group decisions made regarding scope, purpose, and definitions. Members should focus on identifying sections of the text in need of revision, clarification, or documentation, but should not be concerned about the nuances of wording. There will be ample opportunity in future e-mail exchanges to word-smith the final document.

An efficient method of group review of guideline text is to project the relevant section on a screen using a projector and laptop computer. The group moves sequentially through paragraphs of text while the chair, or his/her designee, makes corrections in real time. If a revision is complex or requires additional research, the need is clearly indicated and can be addressed after the meeting by electronic exchange.

### Creating a Draft Guideline From the Topic List

The group now begins the important task of creating a draft list of key action statements for the guideline based on the prioritized topic list. This is not an exercise in linguistic perfection, but instead emphasizes a logical sequence of proposed actions that capture opportunities for quality improvement.

Keeping in mind the quality-driven nature of the process, the chair leads the group in deciding which topics on the prioritized list will be used to create action statements. Invariably there will be some obvious choices that stand out, since they are likely the reason the guideline was undertaken.

Don't be surprised if the task of creating key statements seems awkward or chaotic at first; developing simple, but insightful, action-oriented statements is difficult, even for seasoned guideline writers.

Some considerations in selecting topics include:

1. Try to identify a few critical, "must-have" topics or statements from the list. For example, if the group could change clinician behavior or public policy in only few critical areas, what would they be? If there is a percep-

tion that the condition is currently managed poorly, what should be done differently? Opinions may differ considerably, but some consistent items will invariably emerge.

2. Next, consider topics with high priority scores *and* supporting evidence that includes systematic reviews, randomized trials, or both. Strong evidence supports strong recommendations when the benefit-to-harm balance is favorable.
3. Include topics that can most influence quality of care, *even if* systematic reviews and randomized trials are not available. Recommendations may still be possible if benefits exceed harms, based on lower-quality evidence or working group consensus. Guideline users are often most perplexed by clinical decisions that are not informed by high-quality evidence. Expert consensus about these questions can be quite useful, as long as it is clear that consensus-based recommendations differ from evidence-based recommendations.
4. Develop a logical sequence of statements that moves from broad concepts (eg, diagnosis, modifying factors, initial management) to more specific concerns (eg, therapeutic alternatives, use of diagnostic tests, outcome assessment) to concluding aspects of providing care (eg, treatment failures, recurrence, prevention).
5. Link topics that are deemed not suitable for a key action statement to the supporting text of related statements, thereby allowing the group to include some discussion of the topic without having to justify a distinct action statement and evidence profile.
6. Consider topics with lower priority scores that are nonetheless supported by systematic reviews or randomized trials as potential key action statements, especially when the volume of high-quality evidence is substantial.

As each topic is discussed, the group, assisted by the consultant, should draft a rough version of a related key action statement. This is facilitated with a brief discussion of why the topic was initially proposed and what quality improvement opportunity is offered. The statement should include, as discussed above, the details of "who, what, when, to whom, why, and how."

The most important word in a key action statement is the verb describing the action to be taken. Most guideline-prescribed activities can be described with a limited vocabulary of actions, shown in [Table 7](#).<sup>36</sup> This list is intended not to be restrictive, but rather to help the group in getting started with creating the statements. Other verbs can be used provided they offer clear guidance.

Guideline authors should remember the intended audience and assure that their guidance is applicable. For example, a recommendation that "patients should not be exposed to passive cigarette smoke" is only pertinent if the

**Table 7**  
Suggested action terms for guideline key statements\*

Action	Description
Test	Obtain or collect additional data through inquiry, observation, laboratory testing, or other investigative procedures whose intent is not curative
Prescribe	Order a treatment requiring medication or durable medical equipment
Perform	Perform therapeutic procedure: order activities that are therapeutic in nature
Educate/counsel	Inform the patient about means to improve/maintain health, or instruct on how to perform specific activities
Dispose	Initiate an activity to direct the flow of patients, such as Admit, Discharge, Follow-up, Transfer, etc.
Monitor	Make serial observations according to specific criteria and schedule
Refer/consult	Direct a patient to another clinician for evaluation, treatment, or both
Prepare	Make ready for particular guideline-directed activity by training, equipping, or gaining new knowledge (eg, having procedures in place)
Document	Record one or more facts in the patient record
Advocate	Argue in support of a policy
Diagnose/conclude	Determine a diagnosis or clinical status

\*Listed in order of declining prevalence based on Essaihi et al.<sup>36</sup>

intended audience includes smokers. A preferable statement directed to an intended audience of (non-smoking) clinicians would be to *counsel* patients about the importance of avoiding cigarette smoke.

### Ambiguity and Vagueness in Action Statements

Key action statements should be clear and precise to avoid inconsistent interpretation and prevent inappropriate practice variation. Having drafted a list of key statements, the group should review the list for ambiguous or vague actions.

*Ambiguity* is present when a term can reasonably be interpreted in more than one discrete way.<sup>37</sup> True ambiguity is almost always unintentional and readily correctible when

identified. Examples might include interpretation of an acronym in more than one way (eg, LAD can stand for left anterior descending, left axis deviation, and lymphadenopathy; MS might be interpreted as morphine sulfate, magnesium sulfate, or multiple sclerosis).

- We differentiate truly *ambiguous* statements from *vague* and *underspecified* statements. Vagueness is present when a word's meaning is not well defined, lacking a crisp threshold in a single dimension. Examples include using terms like "short," "febrile," or "old," which are open to broad interpretation. Underspecified statements lack specificity in multiple dimensions, such as "sufficiently ill" or "severe asthma."
- Modifying phrases introduce another form of vagueness. Examples include "it is prudent to recommend." The passive voice is always vague, because the essential "who" of the statement is missing. Similarly, asking clinicians to "consider" an action results in an unmeasurable outcome.
- Sometimes guideline developers introduce statements of fact as recommendations, eg, "adjuvant hormone therapy for locally advanced breast cancer results in improved survival in the long term." Such statements are not executable because the specific action required is never defined.

Although key guideline statements should generally be precise and unambiguous, there may occasionally be a need for deliberate vagueness or underspecification. Reasons for intentionally creating vague recommendations include:<sup>9</sup>

- insufficient evidence: the available literature has not addressed critical topics or the conclusions of published studies are suspect because of methodological flaws
- inability to achieve consensus among the authors regarding evidence quality, anticipated benefits and harms, or interpretation of the science base
- legal considerations: unwillingness to create a potential legal "standard of care"
- economic reasons: one approach is clearly best but may not be affordable
- ethical/religious issues (eg, attitudes about the "the burden" or "futility" of care, premarital sex, use of blood products)

An explicit statement of the reasons for writing deliberately vague recommendations can help users interpret and apply them.

### Refining the Draft Guideline Key Statements

The group should now review and discuss the proposed outline of key action statements for the draft guideline. Is the sequence logical? Have all major quality concerns been addressed? Are there any obvious omissions or inconsistencies? Do the statements reflect the concerns of all disciplines in the working group?

A logical sequence of key action statements is important for clarity and will also facilitate writing if later statements build upon concepts developed in earlier text. The order of statements can always be adjusted later, but effort at this time to ensure smooth, conceptual flow is time well spent.

As mentioned before, the goal is not to have a perfect list of statements, but rather to have an *acceptable* list that serves as a platform for moving forward with guideline development. The list can be revised as the process proceeds, but ideally the agreed-upon structure should be maintained. A draft list of guideline key action statements is shown in Table 8, based upon the topic list for the Hoarseness Guideline shown in the preceding section. Note that the wording is rough and the strength of action (should vs may vs must) will be determined subsequently based on the associated evidence profile at in-person meeting #2.

### Define Supporting Text and Literature Search Needs

Following each key action statement, authors should write several paragraphs of text that develop the rationale for the statement, present the underlying evidence (answering “Why?”), and provide sufficient detail or references so that members of the intended audience will be able to carry it out (answering “How?”). It is important at this stage to remain focused on the recommendation statement and not to strive for encyclopedic coverage of the topic. Concisely stated guidance is more likely to be read and followed.

Although the text will not be written at this time, the group should discuss each key statement sequentially and outline concepts for the supporting text (Table 9). Ideas are recorded under each statement with the results displayed for immediate group feedback. The goal at this stage is to define broad concepts; no attempt is made to create definitive language or statements.

Any topics or issues that were considered important by the group but were not chosen to be key recommendations can nonetheless be discussed in the text under a related key statement heading. By demoting a topic to the amplifying text, information can still be incorporated without the rigor needed to support a key recommendation.

Guideline authors need to be careful not to make substantive recommendations in the amplifying text. In contrast, they should be explicit about the evidence base, or lack thereof, for the actions proposed.

Before moving on to the next key action statement, the group must discuss the need for any stage 3 literature

**Table 8**  
Draft key action statements from the AAO-HNS guideline on hoarseness

1. **Diagnosis:** Clinicians should diagnose hoarseness in a patient with altered vocal quality, pitch, or effort that impairs voice or alters voice-related quality of life.
2. **Laryngoscopy:** The clinician should visualize the larynx of a patient with hoarseness if there is concern of a serious underlying etiology [if possible, list specific criteria here instead of the vague term “concern of a serious underlying etiology].
3. **Modifying Factors:** Clinicians should assess the patient with hoarseness by history and/or physical examination for factors that modify management such as one or more of the following: immunocompromised state, prior laryngeal surgery, [list here, succinctly, all of the major factors; the list does NOT have to be completely inclusive].
4. **Ancillary Testing:** Clinicians should not obtain computed tomography (CT), magnetic resonance imaging (MRI), or electromyography (EMG) of the patient with a primary complaint of hoarseness prior to visualization of the larynx.
5. **Laryngopharyngeal Reflux:** Clinicians may/should/should not routinely prescribe antireflux medications in patients with hoarseness.
6. **Corticosteroid Therapy:** Clinicians should not (routinely) prescribe [oral] corticosteroids to treat patients with hoarseness.
7. **Voice Therapy:** (A) Clinicians should advocate voice therapy for patients diagnosed with hoarseness that persists longer than three weeks or is recurrent and reduces voice-related quality of life. (B) Clinicians should/must visualize the larynx before prescribing voice therapy and document/communicate the results to the speech-language pathologist.
8. Clinicians should not prescribe antibiotics for the treatment of hoarseness or laryngitis in the absence of concurrent bacterial infection.
9. Clinicians should educate patients with hoarseness that surgery is a possible intervention.
10. Clinicians may/should educate/counsel patients with hoarseness about control measures.
11. The clinician may/should prescribe botulinum toxin injections for the treatment of hoarseness spasmodic dysphonia.

searches to support the statement. For example, a search may be needed to fill an evidence “gap” for a key action statement. A statement in the otitis media with effusion guideline reads: “Clinicians should document the laterality, duration of effusion, and presence and severity of associated symptoms at each assessment of the child with otitis media



**Table 9**  
**Sample key action statement with suggestions for writing the supporting text**

**Key action statement:**

Clinicians should use topical antimicrobials for initial therapy of diffuse, uncomplicated acute otitis externa. Systemic antimicrobial therapy should not be used unless there is extension outside the ear canal or the presence of specific host factors that would indicate a need for systemic therapy.

**Concepts to include in the supporting text:**

Define uncomplicated

Explain why oral antibiotics, particularly those approved for children, have limited coverage of acute otitis externa pathogens; conversely, how drops overcome this

Discuss harms: side effects of systemic antimicrobials (allergy, rash, resistance, gastrointestinal upset); how topical drops avoid these problems

Mention that there are no studies of oral vs topical; however, describe cohort studies of outcomes, clinical efficacy, microbiologic efficacy of topical drops

Identify candidates for systemic therapy, in addition to topical therapy: acute otitis media, immunocompromise, diabetes, irradiated bone

Include some discussion or a table of "expert opinion" on practical issues of water restrictions (how long, what method?), showering, using hearing aids, headphones

with effusion."<sup>29</sup> Since no reviews or randomized trials were identified to support this, additional searches were done on the value of documentation, in general, in ambulatory care settings. Two references were identified to support the importance and value of appropriate documentation.

Considerations in planning the stage 3 literature searches include:

- If the key action statement is already supported by systematic reviews or randomized trials, there may be no need for an additional literature search.
- If the key action statement is not supported by high-level evidence, or if additional information is required to answer questions posed in the supporting text, a stage 3 search should be planned. The group specifies the search parameters using a PICO format, which allows the staff lead to conduct an efficient search based on a predefined patient population, intervention, comparison, and outcomes.
- Although the stage 3 searches do not require the methodological rigor used in the stage 1 and stage 2 searches, the staff lead must reasonably document the search terms and processes to ensure transparency in reporting. Addi-

tional internal or external staff may be needed to complete the searches in a timely fashion based on the number required.

- One or more stage 3 searches may be required for the supporting text of a single key action statement. Parsimony and precision are encouraged to limit the search burden and keep the process as focused as possible.

When a key statement is supported by multiple randomized controlled trials, the group should identify and assess existing systematic reviews or meta-analyses. If none exist, an internal systematic review may be planned provided that time, resources, and expertise are available. If an existing, but outdated, systematic review is found, it may be possible to modify or update the data without conducting an entirely new review.

### Writing Assignments and Deadlines

At the conclusion of the first in-person meeting, the chair distributes writing assignments to group members. Group members set and agree to deadlines for preparation of their assignments that will allow the project to remain on track. A democratic process helps to assure adherence.

Each key statement and supporting text is assigned a primary author, who composes a draft that is reviewed by a secondary author with complementary expertise. For example, a statement about surgery might have a surgeon and primary care clinician as primary and secondary authors, respectively. If the group includes representatives of consumer groups and advance practice nursing, they may serve as reviewers, authors, or both. A sample writing assignment grid is shown in Table 10.

The goal of each writing assignment is to ensure that rationale for the related key action statement is explained fully, the logic behind the statement is apparent, all medical terms and actions are clear and unambiguous, and best evidence supporting the statement is explained and referenced. Since most working group members will likely never have encountered this type of assignment, the chair must provide clear and specific written instructions to guide the process (Table 11). The chair should review these instructions with the group to clarify any uncertainty and to emphasize the importance of following them explicitly.

### Key Points to Remember

1. The prioritized topic list, based on e-mail exchange after conference call #2, must be completed before the meeting.
2. The primary goal is to draft the guideline's key action statements and to outline broad concepts that will form the supporting text.
3. Emphasis is placed on creating a clear and logical draft that addresses quality opportunities from the topic list and takes advantage of available high-quality evidence.
4. Writing assignments developed during this meeting will require more time from working group members than any other part of the guideline process.

**Table 10**  
**Sample writing assignment grid from the AAO-HNS hoarseness guideline**

Key statement	Primary author	Secondary author	Reviewer
1. Diagnosis	Member A	Member B	
2. Laryngoscopy	Member C	Member D	
3. Modifying factors	Member E	Member F	
4. Ancillary testing	Member G	Member H	
5. Laryngopharyngeal reflux	Member B	Member D	Member I
6. Corticosteroid therapy	Member F	Member J	Member I
7. Voice therapy	Members J and K	Member L	
8. Antimicrobial therapy	Member M	Member A	Member N
9. Surgery	Member O	Member E	Member G
10. Prevention	Member L	Member O	Member N
11. Botox therapy	Member P	Member C	

## STAGE 3 LITERATURE SEARCH

### Purpose

The final step in identifying evidence is driven by the specific key action statements developed by the working group, which form the core of the guideline. These statements reflect opportunities for education and quality improvement, which may not necessarily be supported by existing systematic reviews or randomized controlled trials. Therefore, the stage 3 searches focus on identifying best published evidence to facilitate writing assignments for specific action statements, and to subsequently assist the group

in determining the corresponding evidence profiles and strengths of recommendation.

As discussed in the preceding section, the group should identify the need for stage 3 searches when discussing the guideline topic list and composing the list of concepts to be covered in the corresponding supporting text.

The stage 3 searches should be coordinated by the staff lead after the first working group meeting. Since multiple

**Table 11**  
**Instructions for group members on composing and completing writing assignments**

*Do not* write a chapter or review article; your contribution should be a clear, logical, focused explanation of why the action statement is important and how it should be implemented. Length should be modest; about 4 to 8 medium-sized paragraphs will be appropriate. If more text is required, use two or more subheadings for organizational clarity.

*First* state why the particular key action statement is important and offer some perspective on the rationale. Next *define* words, phrases, or actions in the key action statement with enough specificity that the reader will understand exactly what is meant or should be done.

*Justify* the key statement with a few paragraphs summarized the supporting literature. Use the highest-quality published evidence, preferably randomized trials, controlled studies, systematic reviews, meta-analyses, or evidence reports.

If parts of the key statement cannot be supported with systematic reviews or randomized trials identified in the stage 1 and 2 searches, promptly contact the chair and staff liaison to arrange for a stage 3 search to fill in gaps. If you identify and use literature not in the original searches, please record the origin and method of identification.

*Literature citations* should be listed in the text as: first author last name and publication year (eg, Smith 2008). If more than one citation exists for that year, distinguish with an "a, b, c," etc. after the year (eg, Smith 2008a, Smith 2008b). At the end of the document, list complete reference citations alphabetically by author last name.

*Discuss* any potential risks, harms, and costs related to the particular recommendation. Support, to the extent possible, by references. State if there is a clear preponderance of benefit or risk, or if risks and benefits appear balanced. Consider the need for *implementation* materials to support the recommendation, such as brochures, teaching aids, etc. The guideline will contain a section on implementation needs, and information provided here will be incorporated.

Please end with any suggestions for *future research* on this topic based on evidence gaps uncovered as you reviewed the supporting literature. The guideline will contain a section on research needs, and information provided here will be incorporated.

searches may be needed, additional staff may be required. Search results should be grouped by major subheadings (eg, etiology, diagnosis, therapy, prognosis) available in standard electronic databases. Search results are sent to the group member assigned as primary writer for the specific action statement. The writer eliminates irrelevant items, leaving a core of evidence that is distributed to the group for reference along with the statement that prompted the literature search.

### Identifying Supplementary Evidence

Supplementary evidence can be identified by using PICO-type questions, which pose a well-focused question in terms of the patient, intervention, comparison, and outcome.<sup>8,38</sup> As an example, consider the question, “In adults with uncomplicated acute bacterial rhinosinusitis, would therapy with amoxicillin, compared with placebo or no medication, improve clinical symptoms in 7 to 10 days?” PICO-type questions facilitate literature searches and can be formulated as follows:

- (P)ATIENT OR PROBLEM: How would you describe the patient population or problem of interest? This may include the disease, surgical procedure, primary problem, coexisting conditions, or demographic factors. Example: “In adults with uncomplicated, acute bacterial rhinosinusitis . . .”
- (I)NTERVENTION: Which main intervention, prognostic factor, or exposure is of interest? Example: “. . . would therapy with amoxicillin . . .”
- (C)OMPARISON: What is the main alternative to compare with the intervention? Example: “. . . compared with placebo or no medication . . .”
- (O)UTCOME: What do you hope the intervention will accomplish, measure, improve, or affect? Example: “. . . improve clinical symptoms in 7 to 10 days?”

If a specific action statement is already supported by high-quality evidence such as systematic reviews or randomized controlled trials, additional stage 3 searches may be unnecessary. When high-quality evidence is sparse, however, a PICO-type question may be formulated for each action statement to facilitate the search.

## WRITING ASSIGNMENTS

### Working Group Assignments

Timely completion of working group assignments is mandatory to keep guideline development on schedule. The chair, staff lead, or both should send reminders in advance of deadlines and monitor when assignments are completed. Delinquent group members should be contacted to ascertain the reason for delay and identify remedial action. In some cases assignments may need to be modified.

### Additional Literature Search and Systematic Review

The need for a search may only become apparent once the author assigned to write the text begins the task and be-

comes more familiar with the material. In this circumstance the author may conduct his or her own search, with or without assistance from the staff lead.

Authors must take care with supplementary literature searches to avoid introducing bias. Search and filter criteria should be documented, regardless of whether the search is conducted by the author, staff lead, or both.

If any systematic reviews or meta-analyses are required, they are planned with the group consultant or methodologist. This will entail substantial added effort because of the methodological rigor needed to produce a valid, unbiased, publication-quality systematic review. Whenever possible, guideline topics should be selected based on a foundation of existing systematic reviews, rather than attempting to create reviews from scratch. Sometimes, however, the need to create a review is unavoidable, and the project should be approached with realistic expectations regarding the time and effort involved.

Systematic reviews or meta-analyses conducted to support the guideline are conducted with an a priori protocol to justify publication as an independent manuscript. Attempting to stuff the review into the guideline text is ill advised, because the reporting detail needed to demonstrate validity of the review will overwhelm the guideline with unnecessary technical details. Only a brief, readable summary is included, with external reference to the systematic review manuscript, which is submitted separately for publication.

### Chair Collates Writing Assignments into First Draft

As the chair receives the completed writing assignments (by electronic mail), they are collated into a first draft of the guideline, using the previously described template (Table 5). The goal is to create the draft before the next in-person meeting, while still allowing sufficient time for preliminary feedback by the group.

In producing the first draft, the chair should strive to maintain the ideas and concepts provided by the writing assignments, although editing may be necessary to standardize the writing style. Writing assignments are best viewed as raw material for composing the draft to ensure a balanced, multidisciplinary product. Authors of the original writing assignment will have ample opportunity to comment on any changes made by the chair in style or format for consistency with the remainder of the guideline.

The effort required by the chair in collating the writing assignments should not be under-estimated. The assignments are likely to

vary greatly in quality, completeness, and punctuality. Consequently, the chair may need to rewrite substantial portions of submissions and fill in conceptual gaps not covered.

When complete, the first draft of the guideline is distributed by e-mail to all working group members for review and comment. Using the “line numbers” feature of the word processor facilitates commenting by line number. The document is distributed as a read-only file (eg, pdf format) to prevent direct modification and force the use of comments based on stable line numbers.

Comments from the group are collected and collated by the staff lead, who then distributes the final list to the chair for incorporation into the draft. Strict accounting of responses to substantive comments is critical to avoid a situation in which the same topic is revisited on multiple occasions. A recommended approach is for the staff lead to create a four-column table (Table 12) in which all comments received from group members are listed by line number. The “disposition” states how the chair handled the comment, and will ensure members that their concerns have been addressed.

The chair revises the guideline draft based on the collated comments and dispositions. Any changes made to the draft are performed using the “track changes” feature of the word processor, so they are easily identifiable. The revised guideline and the summary table of

comments are distributed to working group members before the next meeting.

## UNDERSTANDING EVIDENCE PROFILES

### Evidence Profiles and Transparency

An essential of guideline development is transparency in how policy statements are developed and classified as recommendations. An elegant way of accomplishing this is to add an “evidence profile” after the supporting text for each key action statement that lists all decisions made by the group.

The best time to complete the profile is immediately after the group discusses a specific action statement and the associated supporting text.

Evidence profiles assist guideline writers and users by:

1. encouraging an explicit and transparent approach to guideline writing
2. forcing guideline developers to discuss and document the decision-making process
3. creating “organizational memory” to avoid re-discussing already-agreed-upon issues
4. allowing guideline users to rapidly understand how and why statements were developed

**Table 12**

**Summary table with sample comments received for the first draft of the AAO-HNS otitis media with effusion guideline with disposition**

Line	Source	Comment	Disposition
515 524-26	Fred Sara	Change “may be” to “is” [very important] The statement regarding susceptibility to deficits in older children is not validated. I would delete	Done I agree; the paragraph reads better without it
526	John	. . . reading ability <i>due to prolonged MEE? Or to the hearing loss associated with it?</i>	Sentence was deleted
529	Sara	“Given enough time” we’re all dead! Get rid of that phrase!	Done
533-6	Mike	Would be useful to add the relative risk for these—would indicate the degree of risk they convey	The odds ratios for all are about 2.0-3.0; we decided previously as a group not to put these in
547	Fred	Would start new paragraph at “Conditions”	Done
554 566	Karen John	Define sound field. Not acceptable, “or surgery” should not be there.	Reworded and defined Sorry, but the rest of the group disagrees. We are listing surgery here as an option based on individual circumstances, not as a recommendation as in the prior guideline. The existing evidence is not strong enough to eliminate surgery as an option.



**Table 13**  
Evidence profile constructs for key action statements

Construct	What to include in the profile	Comments
Aggregate evidence quality	Specify as A, B, C, D, or X (see Table 14) based on individual study results, magnitude of the effects, and the individual and aggregate sample sizes of the studies.	There are no strict rules and the level of evidence does not automatically drop to the lowest study type included. Rather, the group should reach a consensus rating and document the rationale.
Benefit	List the favorable changes in outcomes, as defined by the group, that would likely occur if the action statement were followed.	Include qualitative and quantitative information, the latter often abstracted from randomized trials and reviews. Be explicit and comprehensive.
Harm	List the adverse events or other unfavorable outcomes that may occur if the action statement were followed.	Include qualitative and quantitative information, and report with the same rigor and detail used in defining benefits.
Cost	State costs related to the prescribed action.	Include direct and indirect costs.
Benefit-harm assessment	Classify as a “preponderance of benefit over harm” (or vice versa) or a “balance of benefit and harm.”	Stronger recommendations are possible when clear benefit is not offset by important harms or costs (or vice versa); conversely, when the benefit is small or offset by important adverse events, the balance between benefit and harm prevents a strong recommendation.
Value judgments	Summarize value judgments used by the group in creating the action statement; if none were involved, state “none.”	Translating evidence into action often involves value judgments, which include guiding principles, ethical considerations, or other beliefs and priorities; stating them clearly helps users understand their influence on interpreting objective evidence.
Intentional vagueness	State reasons for any intentional vagueness in the action statement; if none was intended, state “none.”	Action statements should be clear and specific, but there may be reasons the group chooses to be vague (eg, concern over setting a legal precedent); acknowledging these clearly promotes transparency.
Role of patient preferences	Specify as large, moderate, small, or none, based on the opportunity for shared decision-making with the patient or proxy.	Weaker evidence with favorable natural history suggests a large role, whereas strong evidence with clear benefit limits the role.
Exclusions	List situations or circumstances where the action statement should not be applied.	Clear exclusions are of particular importance when guidelines are adapted to measuring performance.

#### 5. helping identify aspects of a guideline best suited to performance assessment

Evidence profiles appear immediately after the supporting text for a specific key action statement as a bulleted list with the following headings, defined in Table 13:

- AGGREGATE EVIDENCE QUALITY:
- BENEFIT:
- HARM:
- COST:
- BENEFIT-HARM ASSESSMENT:
- VALUE JUDGMENTS:
- INTENTIONAL VAGUENESS:
- ROLE OF PATIENT PREFERENCES:
- EXCLUSIONS:

The evidence profile makes explicit and transparent the process by which evidence and opinion are transformed into recommendations about appropriate care. The group sequentially describes each aspect of the profile.

To illustrate the structure of evidence profiles, examples from the AAO-HNS guideline on benign paroxysmal positional vertigo are provided below.<sup>6</sup> Each profile lists the associated key action statement followed by the supporting rationale. Details on determining the aggregate evidence quality are provided in the next section.

The first sample evidence profile accompanies a key action statement deemed a “recommendation” in the final guideline based

on evidence quality and harm-benefit assessment. Note the detailed information provided under aggregate evidence quality and exclusions.

**Sample key action statement #1: Clinicians should treat patients with posterior semicircular canal benign paroxysmal positional vertigo with a particle repositioning maneuver. Evidence profile:**

- AGGREGATE EVIDENCE QUALITY: Grade B, based on randomized trials with small sample sizes and significant heterogeneity; most studies were conducted primarily in specialty practice settings with limited data from other treatment settings, potentially limiting generalizability of results
- BENEFIT: Prompt resolution of symptoms with a relatively low number needed to treat ranging from 1 to 3
- HARM: Transient provocation of benign paroxysmal positional vertigo symptoms by the maneuver; risk of falls and imbalance after the procedure; no serious events reported in randomized trials
- COST: Procedural cost
- BENEFIT-HARM ASSESSMENT: Preponderance of benefit over harm
- VALUE JUDGMENTS: High value ascribed to prompt resolution of symptoms and the ease with which the canal repositioning maneuver may be performed
- INTENTIONAL VAGUENESS: None
- ROLE OF PATIENT PREFERENCES: Limited
- EXCLUSIONS: Patients with physical limitations including cervical stenosis, Down syndrome, severe rheumatoid arthritis, cervical radiculopathies, Paget's disease, morbid obesity, ankylosing spondylitis, low back dysfunction, retinal detachment, and spinal cord injuries may not be candidates for this maneuver or may need specialized examination tables for performance of the maneuver

The second sample evidence profile accompanies a key action statement deemed an "option" in the final guideline. Note the relative balance of harm vs benefits and the explicit statements about value judgments and intentional vagueness.

**Sample key action statement #2: Clinicians may offer observation as initial management for patients with benign paroxysmal positional vertigo and assurance of follow-up. Evidence profile:**

- AGGREGATE EVIDENCE QUALITY: Grade C, based on control groups from randomized trials and observational studies with heterogeneity in follow-up and outcome measures
- BENEFIT: Symptom resolution in 15 to 85 percent of patients at one month without intervention

- HARM: Prolonged symptoms compared with other interventions that may expose patients to increased risks for falls or lost days of work
- COST: Indirect cost of delayed resolution
- BENEFIT-HARM ASSESSMENT: Relative balance of benefit and harm
- VALUE JUDGMENTS: Bias of the group for intervention rather than observation, particularly with respect to faster resolution; older patients and those with preexisting balance disorders or high risk for falls may not be suitable for observation
- INTENTIONAL VAGUENESS: "May" is used to allow flexibility in decisions based on the limited evidence and balance of benefit and harm
- ROLE OF PATIENT PREFERENCES: Substantial role for shared decision making
- EXCLUSIONS: None

The third sample evidence profile accompanies a key action statement deemed a "recommendation against" in the final guideline. Note the preponderance of benefit over harm and the exclusions.

**Sample key action statement #3: Clinicians should not routinely treat benign paroxysmal positional vertigo with vestibular suppressant medications such as antihistamines or benzodiazepines. Evidence profile:**

- AGGREGATE EVIDENCE QUALITY: Grade C, based on observational and cross-sectional studies showing uncertain benefits of suppressants in patients with benign paroxysmal positional vertigo
- BENEFIT: Avoid adverse effects, including drug interactions; prevent decreased diagnostic sensitivity during Dix-Hallpike maneuvers
- HARM: Small risk of withholding therapy that may be later shown effective
- COST: None
- BENEFIT-HARM ASSESSMENT: Preponderance of benefit over harm
- VALUE JUDGMENTS: High value ascribed to avoidance of harm from ineffective treatments
- INTENTIONAL VAGUENESS: "Routine" is used since there may be occasional indications for therapy, as defined in the text
- ROLE OF PATIENT PREFERENCES: Minimal
- EXCLUSIONS: Severely symptomatic patients refusing other treatments and patients requiring prophylaxis for canal repositioning maneuvers

**Assigning an Aggregate Level of Evidence**

Guideline groups often have difficulty reaching consensus on the aggregate level of evidence supporting a key action statement. The aggregate is not based simply on the highest- or lowest-quality *single* study identified, but rather a com-

**Table 14**  
**Evidence quality rating scale for grades of evidence**

Grade	Aggregate quality	Comment
A	Well-designed randomized controlled trials or diagnostic studies performed on a population similar to the guideline's target population	Randomized trials should have a valid randomization scheme, blinded (masked) allocation, and minimal loss to follow-up (less than 20%). Similarly, a quality diagnostic test assessment would have clear disease definition and a gold-standard reference criterion. The studies must also be relevant to the key action statement parameters of what, when, to whom, by whom, etc.
B	Randomized controlled trials or diagnostic studies with minor limitations or highly consistent evidence from observational studies	Limitations include nonfatal flaws in methodology, such as incomplete allocation concealment, single or no blinding, limited or incomplete follow-up, or inadequate description of a diagnostic test or reference standard against which the test is compared. Limitations may also be caused by studying a patient population that differs from the target population, resulting in a need for the group to extrapolate findings. Controlled observational studies (eg, case-control or cohort design) can support a grade B assignment if they are well designed, consistent, and relevant to the target population.
C	Observational studies with heterogeneity or with limitations in methodology	Controlled studies are preferred, but for some situations (eg, uncommon disorders or certain procedures) case series of adequate sample size in relevant populations may be used.
D	Expert opinion, case reports, reasoning from first principles (bench research or animal studies)	When the literature search has found limited or anecdotal evidence to support the key action statement the group should consider using consensus to identify current best practice. The process should be robust, thoroughly documented, and based upon group discussion with all stakeholders.
X	Exceptional situations	See text for description.

Adapted from the American Academy of Pediatrics.<sup>16</sup>

posite rating of the quality, consistency, and relevance of the *overall* group of studies. Assigning a rating always incorporates some component of judgment, which is permissible *provided that* the group is consistent and clearly states the reasoning involved under the “aggregate evidence quality” section of the evidence profile (Table 14).

The goal of the evidence review is to determine our *confidence* in the factors of benefit (eg, magnitude of each beneficial effect) offset by our confidence in our understanding of the risks, harms, and costs. Evidence for harms should be assessed with the same diligence applied to evidence for benefit. The purpose of the evidence review is to help us understand the benefit-risk equation.

Many rating scales have been developed, both for individual studies and for aggregate assessments. The scale used by the American Academy of Pediatrics<sup>16</sup> (Table 14) is recommended for clarity and simplicity. A useful scale has also been developed by the United States Preventive Service

Task Force, rating the overall evidence for a service as good, fair, or poor based on study number, quality, consistency, and generalizability.<sup>39</sup>

In nearly all situations the aggregate evidence level can be designated as A, B, C, or D using the criteria in Table 14. Recall, however, that these designations apply *only* to the aggregate evidence level and not to the individual, contributing studies. At times the group may decide to extrapolate evidence from similar, but not directly comparable, patients (eg, using findings from a study of adults in making a recommendation for the pediatric population), especially if high-quality evidence is found.<sup>20</sup> The basis for extrapolating data and the assumptions made should be stated concisely and explicitly as part of the aggregate evidence level description.

A final category of evidence is “Grade X,” used for exceptional situations where validating studies cannot be performed and there appears to be a clear preponderance of benefit or harm. This special category is appropriate in *rare* circumstances where the group has defined a need for a key action statement to improve quality, but the nature of the situation is unlikely to ever result in high-quality evidence.

**Figure 1** A key action statement may be classified as an option, recommendation, or strong recommendation based on the aggregate evidence quality and benefit-harm assessment as determined in the evidence profile. RCT indicates randomized controlled trial. Adapted from the American Academy of Pediatrics.<sup>16</sup>

For example, randomized trials would be unethical to study antimicrobial prophylaxis for anthrax,<sup>16</sup> ototoxic vs non-ototoxic ear drops for acute otitis externa with a non-intact tympanic membrane perforation,<sup>3</sup> or prolonged observation of otitis media with effusion in children with developmental delays or disorders.<sup>29</sup>

## UNDERSTANDING RECOMMENDATION GRADES

Although many different methods have been proposed for grading recommendation strength, most developers agree that determining the strength of action is distinct from rating the aggregate quality of evidence. High-quality evidence (eg, grade A) does not always justify strong recommendations, and recommendations—or even strong recommendations—may be possible despite lower-quality evidence (eg, grade B, C, or X).<sup>40</sup> The primary modifying factor in this regard is the *benefit-harm assessment*, as defined in the preceding section on evidence profiles.

The method for determining strength of recommendation (Fig 1 and Table 15) developed by the American Academy of Pediatrics is simple, transparent, and clinically relevant.<sup>16</sup> Similar to the GRADE approach,<sup>41</sup> the aggregate evidence

level and benefit-harm assessment are the primary rating determinants. GRADE is more complex, however, and offers only two levels of action strength (“strong recommendation” and “weak recommendation”) in contrast to the three levels from the AAP (“strong recommendation,” “recommendation,” and “option”). The authors’ empiric experience in developing guidelines suggests that three levels supports more flexible decision making and is better accepted by clinicians.

Using three levels of action strength is supported by research into the obligation level conveyed by terms commonly found in clinical practice guidelines.<sup>42</sup> Despite a large number of descriptive terms, the obligation levels cluster into three distinct levels: “must” conveys the highest obligation level, “may” the lowest, and “should” an intermediate level. These terms can be used to strengthen a connection between recommendation language and expected adherence to recommendations. For example, a “strong recommendation” carries an obligation of “must” or “should,” a “recommendation” an obligation of “should,” and an “option” an obligation of “may.” “Should” is the most commonly used term in published guidelines.

The strength of action is best viewed as a relative constraint on clinician behavior. In general, less



frequent variation in practice is expected for a strong recommendation than might be expected for a recommendation. The desire of many authors to make uniformly strong recommendations must be tempered by the reality of the evidence quality and benefit-risk assessment.

Assigning a strength of action to a key statement should be very straightforward if performed *after* the evidence profile is constructed. As shown in Figure 1, however, when there is a preponderance of benefit over harm the group may choose between “recommendation” or “strong recommendation” with level B or X evidence quality. Once a choice is made, the reasons should be stated in the evidence profile, usually under the “values” section.

To understand better how the strength of action is determined by the aggregate quality and benefit-harm assessment, consider the statements below from various AAO-HNS guidelines. In each case Figure 1 can be used to cross-check the link between evidence, harm-benefit, and action strength.

The following key action statements are “strong recommendations,” meaning clinicians should follow this guidance unless a clear and compelling rationale for acting in a contrary manner is present.

- **Clinicians should use pneumatic otoscopy as the primary diagnostic method for otitis media with effusion.** *Strong recommendation* based on grade A aggregate evidence (systematic review of cohort studies in relevant populations) with a preponderance of benefit over harm.<sup>29</sup>
- **The management of diffuse acute otitis externa should include an assessment of pain. The clinician should recommend analgesic treatment based on the severity of pain.** *Strong recommendation* based on grade B aggregate evidence (one randomized trial limited to acute otitis externa and consistent, well-designed randomized trials of analgesics for pain relief in general) with a preponderance of benefit over harm.<sup>3</sup>

**Table 15**  
Determining strength of action for key guideline statements

Statement	Definition	Implication
Strong recommendation	A strong recommendation means the benefits of the recommended approach clearly exceed the harms (or, in the case of a strong negative recommendation, that the harms clearly exceed the benefits) and that the quality of the supporting evidence is excellent (Grade A or B)*. In some clearly identified circumstances, strong recommendations may be made based on lesser evidence when high-quality evidence is impossible to obtain and the anticipated benefits strongly outweigh the harms.	Clinicians should follow a strong recommendation unless a clear and compelling rationale for an alternative approach is present.
Recommendation	A recommendation means the benefits exceed the harms (or, in the case of a negative recommendation, that the harms exceed the benefits), but the quality of evidence is not as strong (Grade B or C)*. In some clearly identified circumstances, recommendations may be made based on lesser evidence when high-quality evidence is impossible to obtain and the anticipated benefits outweigh the harms.	Clinicians should also generally follow a recommendation, but should remain alert to new information and sensitive to patient preferences.
Option	An option means either that the quality of evidence is suspect (Grade D)* or that well-done studies (Grade A, B, or C)* show little clear advantage to one approach vs another.	Clinicians should be flexible in their decision making regarding appropriate practice, although they may set bounds on alternatives; patient preference should have a substantial influencing role.

Adapted from the American Academy of Pediatrics<sup>16</sup>

\*See Table 14 for definitions of evidence grades.

The following key action statements are “recommendations,” meaning clinicians should generally follow this guidance but also should be alert to new information and sensitive to patient preferences.

- **Clinicians should treat the patient with cerumen impaction with an appropriate intervention, which may include one or more of the following: cerumenolytic agents, irrigation, or manual removal other than irrigation.** *Recommendation* based on grade B and C aggregate evidence (randomized controlled trials with limitations and cohort studies) with a preponderance of benefit over harm.<sup>5</sup>
- **When the patient (with acute otitis externa) has a tympanostomy tube or known perforation of the tympanic membrane, the clinician should prescribe a non-ototoxic topical preparation.** *Recommendation* based on grade D and X aggregate evidence (reasoning from first principles, and exceptional situations where validating studies cannot be performed) with a preponderance of benefit over harm.<sup>3</sup>
- **Clinicians should not obtain radiographic imaging, vestibular testing, or both in a patient diagnosed with benign paroxysmal positional vertigo, unless the diagnosis is uncertain or there are additional symptoms unrelated to benign paroxysmal positional vertigo that warrant testing.** *Recommendation against* based on grade C aggregate evidence for vestibular testing (diagnostic studies with limitations in referred patient populations, and observational studies) and radiographic imaging (observational studies) with a preponderance of benefit over harm.<sup>6</sup>

The following key action statements are “options,” which offer clinicians flexibility in their decision making but may set boundaries on alternatives. Patient preference should have a substantial role in influencing clinical decision making.

- **Observation without use of antibiotics is an option for selected adults with uncomplicated acute bacterial rhinosinusitis (ABRS) who have mild illness (mild pain and temperature <38.3°C or 101°F) and assurance of follow-up.** *Option* based on grade B aggregate evidence (randomized controlled trials with heterogeneity in diagnostic criteria and illness severity) and relative balance of harm vs benefit for nonsevere ABRS.<sup>4</sup>
- **The clinician may obtain nasal endoscopy in diagnosing or evaluating a patient with chronic rhinosinusitis or recurrent acute rhinosinusitis.** *Option* based on grade D aggregate evidence (expert opinion) with a preponderance of benefit over harm.<sup>4</sup>

Key action statements that lead to recommendations or strong recommendations are most desirable in guidelines,

but options and no recommendations may also serve an important educational role. Options are helpful in addressing controversial aspects of management, especially when wide practice variation exists but evidence is sparse. A clear, systematic review and interpretation of the evidence, using expert consensus to fill gaps, may be very helpful to clinicians, consumers, and policymakers in facilitating decisions.

## IN-PERSON MEETING #2: GRADING RECOMMENDATIONS

### Purpose

The purpose of this meeting is to polish the key action statements, review supporting text, and assign *evidence profiles* to each action statement. Creating evidence profiles should occupy most of the available time, because it is a critical part of guideline development that is best accomplished with an in-person interchange among group members. The evidence profiles are the main determinant of strength for the associated key action statement, based on the aggregate level of supporting evidence and the benefit-harm profile associated with following the action (Table 13).

The main product of the meeting is a second draft of the guideline that accurately reflects the logic and goals of the working group. In addition, the group incorporates suggestions for implementation and future research into the guideline.

### Predistribute Electronic Materials

Documents should be distributed by e-mail prior to the conference call for review by participants before the meeting. Materials for predistribution include:

- agenda for the conference call
- updated working group contact information grid with conflict of interest disclosures
- draft guideline based on writing assignments and initial group feedback
- any other information on content or methodology identified by working group members that would be helpful in educating the group or stimulating discussion

### Reviewing the Draft Guideline

The main goal of reviewing the draft guideline is to achieve a logical, consistent document that accurately reflects the group intentions, minimizing vagueness and under-specification. The goal *is not* to quibble over semantics, grammar, or sentence structure, all of which waste valuable time and can be done through electronic mail exchange.

The chair, assisted by the consultant, must effectively manage time and the group dynamics during the meeting to facilitate steady

progress and to ensure that a minority of voices do not dominate the session.

Begin by reviewing the guideline front matter, which includes information about purpose, and disease burden. Since this was already reviewed once at the prior meeting, the process should not require substantial time. The document is revised in real time by the chair, or his or her designate, and requests for additional information or fact checking are assigned for completion after the meeting.

*Research and implementation needs.* Reviewing the draft guideline presents an ideal opportunity for identifying research and implementation needs, beyond any already specified in the writing assignment. Throughout the meeting thoughts related to opportunities for effective implementation and future research are recorded:

- **IMPLEMENTATION NEEDS.** One group member, ideally an assistant chair, is assigned by the chair to record implementation needs. Consider what exactly will be needed for the target clinician to effectively and efficiently perform what is requested in the key action statement. Examples include fact sheets, brochures, algorithms, visual aids, and videos.
- **RESEARCH NEEDS.** Another member, ideally an assistant chair, is assigned to record ideas for future research. As the group discusses supporting text evidence, voids become apparent, creating the perfect opportunity to decide what is needed to fill the void.

*Key action statement review.* Next, each key action statement is reviewed by the group along with the supporting text. The chair and consultant lead the discussion, striving for balanced input from the group and efficient use of time. The chair or his or her designate record changes to the guideline in real time, projecting the guideline for all to see and using “track changes” on the word processor to clearly identify what has been altered.

The following sequence is advised when reviewing each action statement and supporting text:

1. Assess the introductory paragraphs of supporting text to determine if they effectively engage the reader by giving an overview of the rationale for the key action statement and how following it will improve quality of care.
2. Read the key action statement to the group and identify any potential words or actions that are vague, under-specified, or unclear.
3. Review all remaining paragraphs for content, sequence, and conceptual clarity. Is the purpose of each paragraph clear? Do the paragraphs follow a logical sequence? Did the writers cover all concepts in the original writing assignment?
4. Consider the need for one or more subheadings in the supporting text if the section is long, or if there are several discrete concepts that would benefit from segregation.
5. Discuss the *evidence profile* for the key action statement, once the group has achieved reasonable consensus on supporting

text structure and content. The evidence profile promotes transparency and consistency in decision making, and documents clearly how strength of action is determined.

6. Return to the key action statement and see if any rewording or clarification is needed based on discussion of the items above and the evidence profile. A *strength of action* can now be assigned (Fig 1) using the aggregate level of evidence and the benefit-harm profile.

Repeat the steps above until all key action statements and supporting text are discussed, and consensus has been reached on composition and structure. Some sections will invariably require more discussion than others, but time must be allowed for adequate discussion of all statements and supporting text. If a section requires significant rewriting, reorganization, or additional citations, the specific needs are recorded as an action item that will be addressed immediately after the meeting by the chair or his or her designee.

### Assigning Evidence Profiles and Recommendation Strength

One of the most important goals of sequentially reviewing the key action statements and accompanying text is to reach consensus on the evidence profiles (Table 13). Although the profile statements should be concise, at this stage it is better to be more verbose and err on the side of over-explanation than risk not being clear on why decisions were made. The profiles can subsequently be edited for brevity and consistency.

Evidence profiles are a primary means of promoting transparency in guideline development, and must be developed with care and consistency. Additional time spent ensuring full consensus on the profiles will facilitate grading recommendation strength.

Once the profile has been agreed upon for a specific key action statement, the group determines strength of recommendation using Figure 1 and Table 15. This should be very straightforward if the evidence profile was fully discussed. At times, however, the strength of recommendation does not “make sense” to one or more group members. When this occurs the evidence profile is reviewed, especially the aggregate evidence quality, to ensure accuracy of the information and that no important evidence was overlooked.

An accurate, explicit evidence profile offers the most compelling argument for the group to accept a recommendation grade that challenges existing biases and preconceptions.

### Outline of Key Action Statement Topics

After the key action statements have been developed and the sequence agreed upon by the group, a table is added to the guideline summarizing the statement topics and strengths.

An example is the table included in the AAO-HNS sinusitis guideline (Table 16). The purpose is to orient readers to the structure and content of the guideline, and to highlight stronger statements for easy identification.

### Consider the Need for an Algorithm

Many guidelines benefit from having one or more clinical algorithms that graphically display decision logic and sequences of activities (Fig 2). Including an algorithm in a guideline can 1) rapidly convey the scope and organization of the guideline; 2) result in faster learning, higher retention, and better compliance by the practice community; and 3) specify appropriate indications for particular management strategies.

Algorithms are most useful when the decision logic of a guideline is complex and the temporal sequence of activities is unclear.

Algorithms can be developed using readily available software (eg, Microsoft Publisher or PowerPoint) or with the assistance of a consultant. When designing an algorithm:<sup>43,44</sup>

- Use rounded rectangles to describe a clinical state at entry or completion of a decision sequence; diamond-shaped or hexagonal decision nodes to indicate branch points leading to alternate pathways; and rectangles to indicate diagnostic and therapeutic actions. Diamonds and rectangles correspond to the key recommendations in the guideline).

- Reduce clutter by maintaining a logical flow within the algorithm, usually down the page with side issues, smaller subpopulations, and problems not covered moving off to the side.
- Insert counseling and decision nodes wherever a significant patient preference–dependent decision must be made. Such nodes will generally be placed before treatment rectangles where patient input plays a key role in decision making.
- Be sure that each question node has at least two exit arrows, one for “yes” and one for “no.” Counseling and decision nodes should have at least two exit arrows, corresponding to each major decision outcome, eg, “yes,” “no”; platelet count “<50,000,” “50,000 to 150,000,” or “>150,000.”
- Avoid crossing arrows if at all possible. When crossing is unavoidable, use a curved “overpass” to make it clear that one arrow is crossing over another.

### Define Implementation and Identify Obstacles

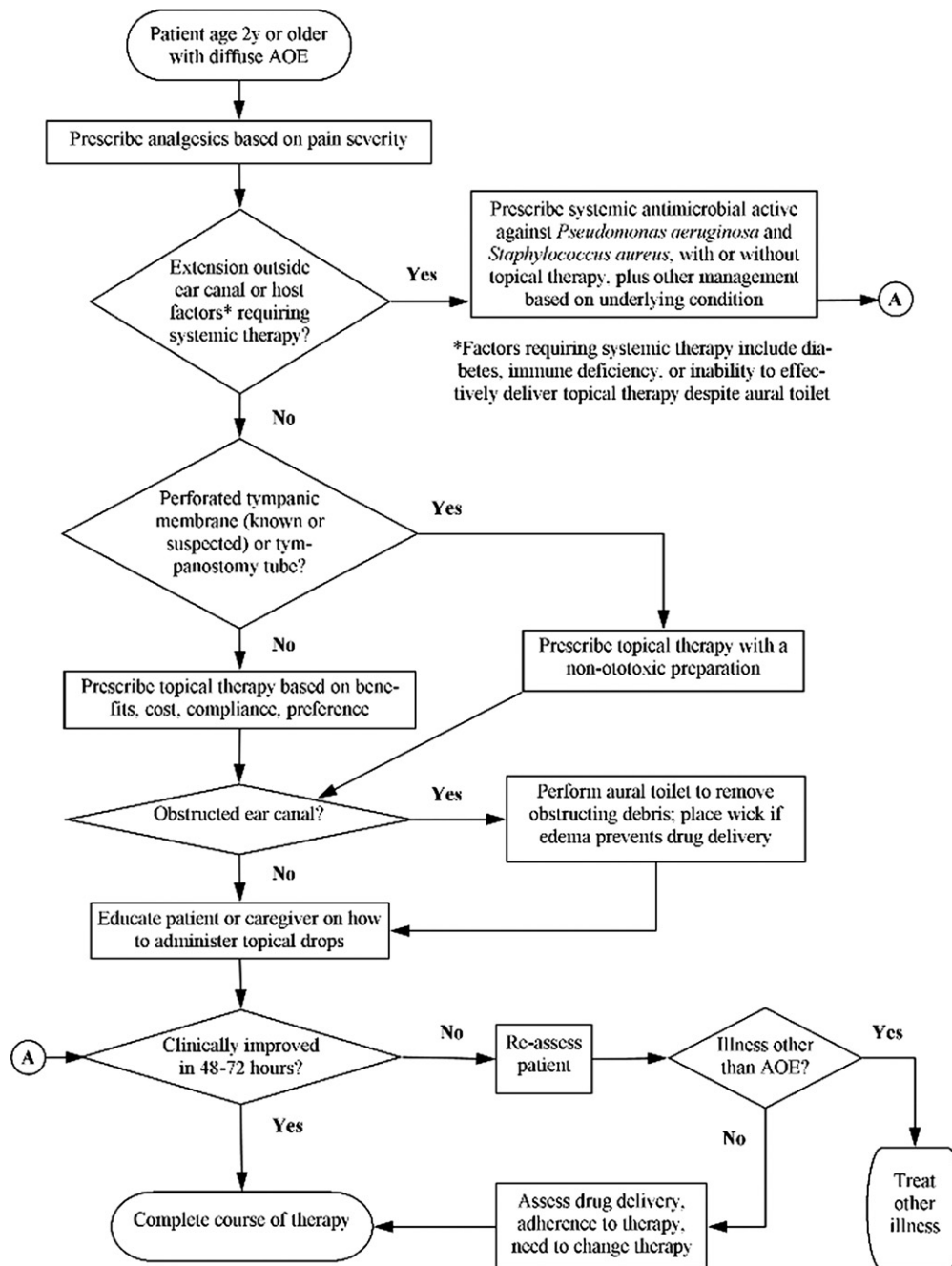
A well-crafted guideline includes a plan for how the recommendations will be implemented, and anticipates obstacles to implementation. As suggested earlier, a running list of implementation issues and needs should be maintained by one of the assistant chairs and updated as discussion proceeds during the second in-person meeting. Issues to consider in this section include 1) plans for distribution and dissemination of the guideline, 2) anticipated obstacles to implementation and proposed solutions, and 3) evaluation

**Table 16**  
Outline of key action statement topics in the AAO-HNS sinusitis guideline<sup>4</sup>

Clinical condition ( <i>key action statement number</i> )	Statement strength*
Presumed viral rhinosinusitis (VRS)	
Diagnosis ( <i>Statement #1a</i> )	Strong recommendation
Radiographic imaging ( <i>Statement #1b</i> )	Recommendation against
Symptomatic relief ( <i>Statement #2</i> )	Option
Presumed acute bacterial rhinosinusitis (ABRS)	
Diagnosis ( <i>Statement #1a</i> )	Strong recommendation
Radiographic imaging ( <i>Statement #1b</i> )	Recommendation against
Initial management	Strong recommendation
Pain assessment ( <i>Statement #3a</i> )	Option
Symptomatic relief ( <i>Statement #3b</i> )	Option
Watchful waiting ( <i>Statement #4</i> )	Recommendation
Antibiotic selection ( <i>Statement #5</i> )	Recommendation
Treatment failure ( <i>Statement #6</i> )	
Subacute sinusitis (no statements)	
Chronic rhinosinusitis (CRS) and recurrent acute rhinosinusitis	
Diagnosis ( <i>Statement #7a</i> )	Recommendation
Modifying factors ( <i>Statement #7b</i> )	Recommendation
Diagnostic testing ( <i>Statement #8a</i> )	Recommendation
Nasal endoscopy ( <i>Statement #8b</i> )	Option
Radiographic imaging ( <i>Statement #8c</i> )	Recommendation
Testing for allergy and immune function ( <i>Statement #8d</i> )	Option
Prevention ( <i>Statement #9</i> )	Recommendation

\*See Table 15 for definitions.





**Figure 2** Example of an algorithm from the AAO-HNS clinical practice guideline on acute otitis externa. Reproduced with permission from the American Academy of Otolaryngology—Head and Neck Surgery.<sup>3</sup>

plans to assess the impact of the guideline on clinical care processes and patient outcomes. The lifespan of the guideline should also be specified, with a statement about when review or revision is planned.

**Working Group Assignments and Next Steps**

The chair reviews specific assignments made during the meeting and assigns deadlines for completion. The chair will also compile a revised guideline based on decisions made at the meeting plus information received from assign-

ments. The final revision should include a section about future research needs.

**Key Points to Remember**

1. The draft guideline, based on e-mail exchange after conference call #2, must be completed before the meeting.
2. The primary goal is to achieve consensus on the organization and content of the guideline.
3. Emphasis is placed on creating clear and logical supporting text that facilitates completing an evidence profile for each

key action statement, which will determine the strength of action for the corresponding key statement.

- Lists are compiled for implementation considerations and future research needs.

## APPRAISING DRAFT GUIDELINE IMPLEMENTABILITY

Guideline appraisal is valuable at this stage to ensure that the guideline is clear, that it adheres to current methodological standards, and (importantly) that recommendations can be *implemented in a manner that is likely to influence clinician behavior*. This can be accomplished by staff at the sponsoring organization, some of whom have not participated in the guideline development process.

The Yale Center for Medical Informatics (YCFMI) has developed a *Guideline Appraisal Report* to aid developers in identifying and remedying potential problems with validity or implementation before publication. Activities involved in the YCFMI guideline appraisal include:

- Marking up the draft guideline to create an XML file using the Guideline Elements Model (GEM).<sup>45</sup> GEM representation is part of a generic process for translating document-based knowledge into workflow-integrated decision tools.
- Assessing the guideline (GEM file) for adherence to the COGS checklist of standardized reporting measures for clinical practice guidelines. The COGS checklist is intended to promote guideline quality by defining components that are necessary for evaluating validity and usability.<sup>14</sup>
- Extracting the decision variables and actions from the guideline recommendation statements and presenting them, free of context, for judgment about decidability and executability.
- Appraising the guideline for implementability using the GuideLine Implementability Appraisal (GLIA), an instrument developed by YCFMI, to assess potential obstacles to guideline implementation.<sup>46</sup> The GLIA addresses eight dimensions of guideline implementability (Table 17).

The following examples illustrate how feedback from the GLIA assessment identifies areas for improvement in guideline statements. In each case the specific key action statement under analysis is stated, followed by the general and specific areas of concern:

**GLIA analysis of draft statement on drug delivery from the acute otitis externa guideline: Clinicians should inform patients how to administer topical drops. When the ear canal is obstructed, delivery of topical antimicrobials should be enhanced by aural toilet, placing a wick, or both.**

1. The recommended action (what to do) is vague or underspecified: some vagueness remains over when to do both aural toilet *and* wick placement.

**Table 17**  
Dimensions of implementability for guideline recommendations

Dimension	Definition
1. Decidability	Precisely under what circumstances to do something
2. Executability	Exactly what to do under the circumstances defined
3. Effect on process of care	The degree to which the recommendation impacts upon the usual workflow in a typical care setting
4. Presentation and formatting	The degree to which the recommendation is easily recognizable and succinct
5. Measurable outcomes	The degree to which the guideline identifies markers or endpoints to track the effects of implementation of this recommendation
6. Apparent validity	The degree to which the recommendation reflects the intent of the developer and the strength of the evidence
7. Novelty and innovation	The degree to which the recommendation proposes behaviors considered unconventional by clinicians or patients
8. Flexibility	The degree to which a recommendation permits interpretation and allows for alternatives in its execution

Adapted from Shiffman et al.<sup>46</sup>

2. The recommendation can not be performed by the guideline's intended users without the acquisition of new competence (knowledge, skills): primary care physicians may not have skills in wick placement/aural toilet.
3. The recommendation does not specify patient or practice characteristics (clinical and nonclinical) that require (or permit) individualization: not clear about when to use which technique.

**GLIA analysis of draft statement on watchful waiting from the acute sinusitis guideline: Observation without use of antibiotics is an option for selected adults with uncomplicated acute bacterial rhinosinusitis based on illness severity and assurance of follow-up.**

1. The recommendation (and its discussion) is not concise: the raw meta-analysis data in the discussion could be referenced rather than included.
2. The recommendation is not compatible with existing attitudes and beliefs of the guideline's intended users: may require convincing clinicians not to use antibiotics.
3. The recommendation does not consider coincident drug therapy and common comorbid conditions: is this rec-

ommendation valid in all populations (eg, elderly, diabetics, hospitalized patients)?

### CONFERENCE CALL #3: ENSURING IMPLEMENTABILITY

The purpose of this call is to review the *Guideline Appraisal Report*, address any deficiencies identified, and plan for external peer review. In advance of the call the *Guideline Appraisal Report* should be distributed to group members.

- The section of the report on adherence to the COGS checklist of standardized reporting measures for clinical practice guidelines is reviewed and discussed. If any deficiencies are noted they are discussed and corrected if necessary.
- The section on guideline implementability (GLIA) is reviewed and discussed sequentially for each key action statement. If any implementation barriers were identified the group discusses potential solutions, which typically involve rewording or clarifying segments of the text.

Now is *not* the time for major changes in the structure or order of key action statements unless the GLIA report identifies a serious deficiency that requires corrective action. The group should focus on remedying barriers to implementation identified in the report, without revisiting peripheral issues. If the conversation does digress, often the evidence profiles can be used as “organizational memory” as to why earlier decisions were made.

Suggestions are next solicited for external peer reviewers to review the draft guideline. Peer reviewers should represent the intended target audience and practice settings, and are selected with input from the chair, working group members, Academy committee chairs, specialty leaders, and others.

External, multidisciplinary peer review is essential to ensure guideline clarity and anticipate concerns or objections before the document is published. Several reviewers are solicited from each discipline involved in the guideline to increase the chance of comprehensive feedback. Typically as many as 25 to 35 reviewers are solicited.

### PRE-RELEASE PEER REVIEW OF DRAFT GUIDELINE

Independent, external peer review of a guideline is a critical aspect of development. Existing guideline processes have been

criticized as biased, self-promoting, and exempt from accepted procedures for scientific publication and editorial peer review.<sup>47</sup> The implementability assessment described in the preceding section promotes clarity of action, but does not address whether the recommended actions are appropriate and meaningful in the first place. This latter concern is the subject of multidisciplinary peer review as described in this section.

A final draft of the guideline is prepared using the “line numbers” feature of the word processor to insert continuous line numbers along the left margin. The draft guideline is distributed electronically to external peer reviewers with instructions to submit comments *by line number*. A strict deadline is specified by which time comments should be submitted to the chair and staff lead.

Distributing the draft guideline as a pdf file has the advantage of making it impossible for a reviewer to simply type changes into the guideline text (which will greatly complicate the chair’s ability to identify them). The pdf file also ensures that all reviewers use the same line numbers, since numbers may not match when different word processing programs or operating systems are used.

Peer reviewers should be asked to focus on three main guideline attributes: validity, reliability, and feasibility:

1. *Valid* guidelines include all relevant literature, have explicit links between decisions and scientific evidence, and clearly distinguish and justify situations where expert judgment or group consensus is used to support recommendations.
2. *Reliable and reproducible* guidelines allow a knowledgeable peer reviewer to arrive at conclusions similar to those of the development group when considering the evidence.
3. *Feasible* guidelines are clearly written, are user friendly, allow for flexibility in individual clinician decisions, and are suitable for routine use in intended settings.

Comments from the external reviewers are collected and collated by the chair. A recommended approach is for the chair to create a four-column table (Table 18) in which all comments received are listed by line number. The “dispositions” state how the chair handled the comment, and will ensure external reviewers that their concerns have been addressed.

The chair revises the guideline draft based on the collated comments and dispositions. Any changes made to the draft are performed using the “track changes” feature of the word processor, so they are easily identifiable. The revised guideline and the summary table of comments are distributed electronically to external peer reviewers.

The final guideline draft and the summary table of external reviewer comments are distributed electronically to the working group for review and approval.

**Table 18**  
**Excerpt of summary table for external reviewer comments from the AAO-HNS guideline on acute otitis externa**

Line	Reviewer	Comment	Disposition
498	Dr. A	Does the use of the term here “antimicrobial” mean both antibiotics and agents like acetic acid? If so, this term could be misconstrued.	The first paragraph below the bold text defines antimicrobial to include an antibiotic, antiseptic, or steroid.
507	Dr. B	Table 6: confirm that Cipro HC is still available	I checked with the manufacturer and it is still available.
579	Dr. C	I noted that the article by Bassim and Drake was left off. It appeared in the ENT Journal this past summer and referred to the predilection for Ciprodex drops to precipitate in the ear, especially on PE tubes and even in the middle ear after going through the lumen of tubes. I really think that this is a point which should be mentioned in the adverse effects section.	The Bassim and Drake article describes 9 cases of delayed tympanostomy tube obstruction after using a steroid-antibiotic drop. None of the cases dealt with acute otitis externa alone. Extrapolating these results to acute otitis externa is beyond the methodological standards we followed for the rest of the guideline.
669	Dr. D	I would be reluctant to encourage irrigations of an acutely infected ear, especially if this recommendation is going to nonspecialist healthcare providers who may have limited experience with aural care. Gentle irrigation at most should be considered. I would favor staying with dry cleaning and suctioning.	Your concern is appreciated, and the wording of the current statement (which includes the word “gentle”) was arrived at after much discussion and consensus and specialists and primary care physicians. To soften the wording, I changed the sentence from “Aural toilet is <i>most often</i> done with a gentle lavage . . .” to “Aural toilet <i>may be</i> done . . .”
692	Dr. B	Personally I think that wicks should only remain in ear 6-8 hours, not be left indwelling for days at a time. I can discuss technique (as listed in the aforementioned book by Lucente).	Changed to: “A wick is unnecessary once the ear canal edema subsides, which may occur within 24 hours (Lucente 1995) or a few days of topical therapy.” I used “24 hours” (instead of 6-8) since that is what your book states several times on pg. 62.

## ORGANIZATIONAL BOARD APPROVAL OF GUIDELINE

Prior to publication, the guideline should be distributed for approval to the Board of Directors of the sponsoring organization(s). The procedure will vary based on organizational policy, but the process used at the AAO-HNS is:

1. The staff lead prepares an Executive Summary of the guideline that contains 1) the Introduction section, 2) the Purpose section, 3) all key action statements with their corresponding evidence profiles, and 4) the Implementation section.
2. The Executive Summary and full-text guideline are distributed to the Board members for review, comment, and approval. The summary grid of external peer review comments and their disposition may also be included.

Because the document is based on evidence, any substantive changes requested by an oversight body (eg, the organizational Board of Directors) must be supported and accompanied by evidence. The oversight body should be informed, however, that the purpose of review is not to rewrite the guideline, but rather to ensure that the recom-

mended actions are consistent with the organization’s mission and values.

Any comments or concerns expressed by the board are responded to by the chair, with input from the working group solicited as needed.

## PUBLICATION AND DISCLAIMERS

The final guideline is converted by staff liaison and chair into a document that meets publication requirements of the sponsoring organization’s official journal.

At the AAO-HNS, the managing editor of the journal is notified that the guideline should not be submitted for external peer review, because it has already been extensively reviewed and further changes are not possible. A copy of the chair’s summary table of external reviewer comments and their disposition should be submitted to the managing editor to document the external review and remain on file in lieu of traditional editorial peer review. Depending on the length of the document, the guideline may be published as a supplement or within the main journal.



If the guideline is lengthy and published as a supplement, an Executive Summary may be prepared for publication in the main journal to promote awareness. The Summary should contain condensed versions of the introductory segments, a tabular listing of all key action statements, full evidence profiles for the action statements, and condensed versions of supporting text.

The guideline page proofs should be checked expeditiously. The guideline chair should prepare a summary of newsworthy points from the guideline, which the organization's public relations department can use in preparing a press release. Embargo and publication dates are coordinated among all involved organizations.

Guidelines should be published with an accompanying disclaimer, created by the organization's legal counsel, to set clear bounds on the intended use of the document. The AAO-HNS adds a brief disclaimer to the abstract and a longer disclaimer to the end of the manuscript. Having a disclaimer in the abstract is advised because individuals without access to the full text may cite only the abstract text.

Here is the AAO-HNS disclaimer text added to the end of the abstract:<sup>6</sup> "This clinical practice guideline is not intended as a sole source of guidance in managing [topic specified here]. Rather, it is designed to assist clinicians by providing an evidence-based framework for decision-making strategies. The guideline is not intended to replace clinical judgment or establish a protocol for all individuals with this condition, and may not provide the only appropriate approach to diagnosing and managing this problem."

Here is the AAO-HNS disclaimer text added to the end of the manuscript:<sup>6</sup> "As medical knowledge expands and technology advances, clinical indicators and guidelines are promoted as conditional and provisional proposals of what is recommended under specific conditions, but they are not absolute. Guidelines are not mandates and do not and should not purport to be a legal standard of care. The responsible physician, in light of all the circumstances presented by the individual patient, must determine the appropriate treatment. Adherence to these guidelines will not ensure successful patient outcomes in every situation. The American Academy of Otolaryngology—Head and Neck Surgery (AAO-HNS) emphasizes that these clinical guidelines should not be deemed to include all proper treatment decisions or methods of care, or to exclude other treatment decisions or methods of care reasonably directed to obtaining the same results."

## IMPLEMENTATION

Publishing the guideline is the first step towards implementation, but publication alone is unlikely to change clinician

behavior. The implementation plan outlined in the guideline is begun with necessary resources committed by the sponsoring organization or external funding agencies. Relevant brochures, educational materials, and continuing medical education (CME) produced by the organization should be updated for consistency with the new guideline recommendations.

Awareness of the guideline is increased by ensuring that the National Guideline Clearinghouse (NGC) receives a copy of the final publication with appropriate copyright release to produce a summary document. Efforts to create performance measures based on the guideline should be planned when appropriate. Such efforts are facilitated by clear action-oriented recommendations that are supported by high-quality evidence.

In general, performance measures are likely to be most valid when built around strong recommendations, where the benefit-risk deliberation shows a preponderance of one or the other and good-quality evidence supports the policy.

Performance measures are *unlikely* to be valuable if built around 1) statements that are vague or underspecified, 2) recommendations where anticipated benefit is balanced by anticipated risk, harm, and cost, or 3) recommendations based on evidence that may change.

## UPDATE

Guidelines should describe how and when the need for an update will be assessed. Situations that might require clinical guidelines to be updated include:<sup>48</sup>

- changes in evidence that bear 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

There are several possible methods for deciding when to update:<sup>48</sup>

1. Include a date for scheduled review in the guideline document. This could waste resources, however, if a full review and update is undertaken prematurely in a slowly changing field or when no significant new evidence is available.
2. Include a statement that an update will occur "when new information becomes available." Although good in theory, there are no validated methods to systematically assess in an ongoing manner exactly when new information becomes available.

3. Ask a multidisciplinary group of experts to review the guideline. Content experts could be recruited from the original guideline group, complemented by others with expertise in critical appraisal. This group would be asked to 1) identify any changes in interventions available, and 2) decide if any new evidence or developments invalidate the guideline's policy statements.
4. Implement a policy stating that guidelines must be reaffirmed, revisited, or retired at a specific time interval (eg, 3 to 5 years), as done by some.

When an update is required, a decision must be made whether to retain and revise the guideline, or to replace it entirely. This decision is necessarily subjective, taking into consideration the breadth and quality of new evidence plus the number of policy statements that are outdated.

## ACKNOWLEDGEMENTS

We thank Jenissa Haidari, MPH, and Miles Patel, MS, for their critical review of the manuscript and helpful suggestions, and David R. Nielsen, MD for his support of our collaborative efforts to create superior guidelines and knowledge products with the AAO-HNSF.

## AUTHOR INFORMATION

From the Department of Otolaryngology, State University of New York Downstate and The Long Island College Hospital, Brooklyn, NY (Dr Rosenfeld); and the Department of Pediatrics and the Yale Center for Medical Informatics, Yale University School of Medicine, New Haven, CT (Dr Shiffman).

Corresponding author: Richard M. Rosenfeld, MD, MPH, Department of Otolaryngology, Long Island College Hospital, 339 Hicks Street, Brooklyn, NY 11201.

E-mail address: [richrosenfeld@msn.com](mailto:richrosenfeld@msn.com).

## AUTHOR CONTRIBUTIONS

Richard Rosenfeld, writer; Richard Shiffman, writer.

## DISCLOSURES

**Competing interests:** None.

**Sponsorships:** None.

## REFERENCES

1. Chinese proverbs quotes. <http://www.thinkexist.com/> (accessed April 12, 2009).
2. Rosenfeld RM, Shiffman RN. Clinical practice guidelines: a manual for developing evidence-based guidelines to facilitate performance measurement and quality improvement. *Otolaryngol Head Neck Surg* 2006;135(Suppl):S1–28.
3. Rosenfeld RM, Brown L, Cannon CR, et al. Acute otitis externa clinical practice guideline. *Otolaryngol Head Neck Surg* 2006;134(Suppl):S4–23.
4. Rosenfeld RM, Andes D, Bhattacharyya N, et al. Clinical practice guideline: adult sinusitis. *Otolaryngol Head Neck Surg* 2007;137(Suppl):S1–31.
5. Roland PS, Smith TL, Schwartz SR, et al. Clinical practice guideline: cerumen impaction. *Otolaryngol Head Neck Surg* 2008;139(Suppl):S1–21.
6. Bhattacharyya N, Baugh RF, Orvidas L, et al. Clinical practice guideline: benign paroxysmal positional vertigo. *Otolaryngol Head Neck Surg* 2008;139(Suppl):S47–81.
7. Schwartz SR, Cohen SM, Dailey SH. Clinical practice guideline: hoarseness. *Otolaryngol Head Neck Surg* 2009 (forthcoming).
8. Eden J, Wheatley B, McNeil B, et al, editors. Knowing what works in health care. A roadmap for the nation. Institute of Medicine of the National Academies. Washington (DC): National Academies Press; 2008.
9. Shiffman RN, Marcuse EK, Moyer VA, et al. Toward transparent clinical policies. *Pediatrics* 2008;121:643–6.
10. Field MJ, Lohr KN, editors. Clinical practice guidelines: Directions for a new program. Institute of Medicine. Washington (DC): National Academies Press; 1990.
11. Turner T, Misso M, Harris C, et al. Development of evidence-based clinical practice guidelines (CPGs): comparing approaches. *Implement Sci* 2008;3:45.
12. Woolf SH, Grol R, Hutchinson A, et al. Potential benefits, limitations, and harms of clinical guidelines. *BMJ* 1999;318:527–30.
13. The AGREE Collaboration. Appraisal of Guidelines for Research & Evaluation (AGREE) Instrument. <http://www.agreecollaboration.org/> (accessed April 14, 2009).
14. Shiffman RN, Shekelle P, Overhage M, et al. Standardized reporting of clinical practice guidelines: A proposal from the conference on guideline standardization. *Ann Intern Med* 2003;139:493–8.
15. National Guideline Clearinghouse. Inclusion criteria. <http://www.ngc.gov/submit/inclusion.aspx> (accessed April 1, 2009).
16. AAP Steering Committee on Quality Improvement and Management. Policy statement: classifying recommendations for clinical practice guidelines. *Pediatrics* 2004;114:874–7.
17. Shekelle PG, Woolf SH, Eccles M, Grimshaw J. Clinical guidelines: developing guidelines. *BMJ* 1999;18:593–6.
18. Guyatt G, Gutterman D, Baumann MH, et al. Grading strength of recommendations and quality of evidence in clinical guidelines. Report from an American College of Chest Physicians Task Force. *Chest* 2006;129:174–81.
19. American Academy of Otolaryngology—Head and Neck Surgery. Quality improvement and patient safety information. <http://www.entnet.org/Practice/quality.cfm> (accessed April 14, 2009).
20. National Institute for Health and Clinical Excellence. The guidelines manual. London: National Institute for Health and Clinical Excellence; April 2007. Available at: [www.nice.org.uk](http://www.nice.org.uk).
21. Scottish Intercollegiate Guidelines Network. SIGN 50. A guideline developer's handbook. Edinburgh: Scottish Intercollegiate Guidelines Network; January 2008. Available at: [www.sign.ac.uk](http://www.sign.ac.uk).
22. U.S. Cochrane Center. Consumers United for Evidence-based Healthcare (CUE). <http://apps1.jhsph.edu/cochrane/uscccc.htm> (accessed January 3, 2009).
23. Tricoci PT, Allen JM, Kramer JM, et al. Scientific evidence underlying the ACC/AHA clinical practice guidelines. *JAMA* 2009;310:831–41.
24. Moher D, Cook DJ, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement. *Lancet* 1999;354:1896–900.
25. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA* 2000;283:2008–12.
26. Rosenfeld RM, Singer M, Wasserman JM, et al. Systematic review of antimicrobial therapy for acute otitis externa. *Otolaryngol Head Neck Surg* 2006;134(Suppl):S24–48.
27. Rosenfeld RM, Singer M, Jones S. Systematic review of antimicrobial therapy in patients with acute rhinosinusitis. *Otolaryngol Head Neck Surg* 2007;137(Suppl):S32–S45.

28. Montori VM, Wilczynski NL, Morgan, et al. Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. *BMJ* 2005;330:68.
29. Rosenfeld RM, Culpepper L, Doyle KJ, et al. Clinical practice guideline: otitis media with effusion. *Otolaryngol Head Neck Surg* 2004; 130:S95–118.
30. Choudhry NK, Stelfox HT, Detsky AS. Relationships between authors of clinical practice guidelines and the pharmaceutical industry. *JAMA* 2002;287:612–7.
31. Higgins JPT, Green S, editors. *Cochrane handbook for systematic reviews of interventions*, version 5.0.1 (updated September 2008). The Cochrane Collaboration; 2008: section 6.4.11.1. Available at: [www.cochrane-handbook.org](http://www.cochrane-handbook.org).
32. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;17:1–12.
33. Jüni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. *BMJ* 2001;323:42–6.
34. Wharam JF, Sulmasy D. Improving the quality of health care: who is responsible for what? *JAMA* 2009;301:215–7.
35. Hussain T, Michael G, Shiffman RN. The Yale Guideline Recommendation Corpus: a representative sample of the knowledge content of guidelines. *Int J Med Inform* 2009;78:354–63.
36. Essaihi A, Michel G, Shiffman RN. Comprehensive categorization of guideline recommendations: creating an action palate for implementers. *AMIA Annu Symp Proc* 2003;220–4.
37. Codish S, Shiffman RN. A model of ambiguity and vagueness in clinical practice guideline recommendations. *AMIA Annu Symp Proc* 2005;146–50.
38. Schardt C, Adams MB, Owens T, et al. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Med Inform Decis Mak* 2007;7:16.
39. U.S. Preventive Services Task Force. Evidence grade definitions. [www.ahrq.gov/clinic/uspstf/grades.htm](http://www.ahrq.gov/clinic/uspstf/grades.htm) (accessed April 12, 2009).
40. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6.
41. Grading of Recommendations, Assessment, Development and Evaluation (GRADE). [www.gradeworkinggroup.org](http://www.gradeworkinggroup.org) (accessed February 23, 2009).
42. Lomotan E, Michel G, Lin Z, et al. How “should” we write guideline recommendations? Interpretation of deontic terminology in clinical practice guidelines: survey of the health services community. *Qual Saf Health Care* 2009 (forthcoming).
43. Society for Medical Decision Making Committee on Standardization of Clinical Algorithms. Proposal for clinical algorithm standards. *Med Decis Making* 1992;12:149–54.
44. Hadorn DC. Use of algorithms in clinical guideline development. In: *Clinical Practice Guideline Development: Methodology Perspectives*. AHCPR Pub. No. 95-0009. Agency for Health Care Policy and Research: Rockville (MD); Jan. 1995. p. 93–104.
45. Shiffman RN, Michel G, Essaihi A, et al. Bridging the guideline implementation gap: a systematic, document-centered approach to guideline implementation. *J Am Med Inform Assoc* 2004;11: 418–26.
46. Shiffman RN, Dixon J, Brandt C, et al. The GuideLine Implementability Appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. *BMC Med Inform Decis Mak* 2005;5:23–31. Available at: [www.biomedcentral.com/1472-6947/5/23](http://www.biomedcentral.com/1472-6947/5/23).
47. Sniderman AD, Furberg CD. Why guideline-making requires reform. *JAMA* 2009;301:429–31.
48. Shekelle P, Eccles MP, Grimshaw M, et al. When should clinical guidelines be updated? *BMJ* 2001;323:155–7.