

CDC Guidelines

Improving the Quality

September 1996



U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Centers for Disease Control and Prevention (CDC)
Atlanta, Georgia 30333



David Satcher, MD, PhD
Director

Dixie E. Snider, Jr., MD, MPH
Associate Director for Science
Office of the Director

Stephen B. Thacker, MD, MSC
Director
Epidemiology Program Office

Steven M. Teutsch, MD, MPH
Chief
Prevention Effectiveness Activity
Office of the Director
Epidemiology Program Office

Centers for Disease Control and Prevention

CDC Guidelines

Improving the Quality

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Centers for Disease Control and Prevention
Epidemiology Program Office
Prevention Effectiveness Activity

Suggested Citation

Centers for Disease Control and Prevention. CDC GUIDELINES: Improving the Quality.
Atlanta: Centers for Disease Control and Prevention, 1996.

Ordering Information

To get bound copies of this guide, call the Prevention Effectiveness Activity at
(404) 639-4455 or write to the following address:

Prevention Effectiveness Activity
Office of the Director
Epidemiology Program Office
Centers for Disease Control and Prevention
Mailstop D-01
1600 Clifton Road, NE
Atlanta, GA 30333

Foreword

As the Nation's prevention agency, the Centers for Disease Control and Prevention (CDC) and the Agency for Toxic Substances Disease Registry (ATSDR) develop, endorse, and disseminate guidelines as a crucial means of improving preventive health care programs and of exercising leadership in public health policy development and implementation. Guidelines translate research findings about the effectiveness and economic impact of prevention programs into accessible and useable information for public health practice. Increasingly, practitioners, the public, and experts in guidelines development methods are demanding practice guidelines that are clear, practical, and based on compelling scientific evidence. CDC must adopt the emerging standards of quality for the guidelines we produce or endorse if it is to maintain its leadership role in public health policy development.

This guide was prepared by a multidisciplinary working group of CDC staff and reviewed by public health and scientific professionals (from federal and state agencies and the private sector) with skill and experience in planning and developing guidelines. They synthesized the best available advice in order to offer CDC staff a set of recommendations and "points to consider" in the planning and execution of each of 13 primary tasks associated with the guideline development process. I encourage you to use this document to train CDC staff and participants in the guideline development process, establish a common language and frame of reference for planning and developing guidelines, and pursue research priorities in the area of guidelines development methods. In addition, I have asked our Centers, Institutes, and Offices to implement the recommendations. I look forward to your full cooperation as we put in place a vigorous regimen of training, technical assistance, and monitoring of results agency-wide. Together we must continually ensure that the guidelines we develop or endorse are science-based, clear and practical. This reaffirms the agency's commitment to excellence.



David Satcher, MD, PhD

Director

Centers for Disease Control and Prevention, and

Administrator

Agency for Toxic Substances Disease Registry

Contributors

Guidelines Methodology Working Group (GMWG)

CDC GUIDELINES: Improving the Quality was completed under the direction of the GMWG which was convened as an ad hoc subcommittee of the Excellence in Science Committee, Office of the Associate Director for Science, Office of the Director, CDC. The members of the GMWG were:

Dixie E. Snider, Jr. M.D., CDC/OD — Co-Chairman
Steven M. Teutsch, M.D. EPO/OD — Co-Chairman
Benedict I. Truman, M.D. EPO/OD — Coordinator/Scientific Editor
Robert E. Keaton, B.A., — Writer/Editor (Contractor)

ATSDR:	Charlie Xintaras, ScD
CDC/OD/MCC:	Dan Sosin, MD
NCCDPHP:	Suzanne Smith, MD
NCEH:	Peter Briss, MD
	Andrew Davis, MD
NCHSTP:	Phil Nieburg, MD
	Samuel Dooley, MD
NCID/HIP:	Ofelia Tablan, MD
NCIPC:	Jim Belloni, MA
NIOSH:	Linda Martin, PhD
NIP:	Stephen Hadler, MD
PHPPPO:	Ray (Bud) Nicola, MD
	Randy Gordon, MD

Authors (Listed Alphabetically)

Andrew M. Davis, MD, MPH
Diane Dunet, MPA
Robert Keaton, BA
Dixie E. Snider, Jr. MD, MPH
Daniel M. Sosin, MD, MPH
Donna F. Stroup, PhD, MSc
Steven M. Teutsch, MD, MPH
Stephen B. Thacker, MD, MSc
Benedict I. Truman, MD, MPH

Acknowledgments

In addition, we wish to acknowledge the valuable assistance of the following persons:

Editing and Editorial Design

Robert Keaton, BA, Writer-Editor

Reviewers Who Made Comments

Richard Ehrenberg, MD
Marilyn J. Field, PhD (IOM)
Richard A. Goodman, MD, MPH
Suzanne M. Hewitt, MPA.
Alan Hinman, MD, MPH
Wanda Jones, Dr PH
Douglas B. Kamerow, MD, MPH (AHCPR)
Kevin Malone, JD
Verla Neslund, JD
Stephen Ostroff, MD
Elaine Power, MPP (FDA)
Deborah Tress Weimer, JD

Production and Technical Support

Sheila Allgood
Martha Boyd (graphics)
Beverly J. Holland (graphics)
Peter Jenkins (layout and typesetting)
Mary Moreman
Charles H. Yancey (printing)

Table of Contents

Foreword	iii
Contributors	iv
Guidelines Methodology Working Group (GMWG)	iv
Authors (Listed Alphabetically)	iv
Reviewers Who Made Comments	v
Executive Summary	1
Introduction	11
Purpose and Utility	11
Mandate and Process	11
Audience and Scope	12
Terminology and Usage	13
Background	15
Assumptions	15
Public Health Practice at CDC	15
Public Health Practice Guidelines	16
Rationale for Using Guidelines in Public Health Practice	17
Desirable Attributes of Guidelines	17
Guidelines Development at CDC	18
Improving the Quality of CDC Guidelines	19
Ethical Issues	20
Controversy and Consensus	20
Planning and Coordinating the Process	21
Points to Consider	21
Recommendations	21
Approaches to Planning and Producing Guidelines	22
Administrative Issues	24
Assessing User Needs	27
Points to Consider	28
Recommendations	28
Approaches to Assessing User Needs	29
Choosing Guideline Topics	33
Points to Consider	34
Recommendations	35
Selecting Guideline Panels	37
Points to Consider	37
Recommendations	38
Approaches to Selecting Panels	39
Current Practice at CDC	41
Defining the Scope of a Guideline	43
Points to Consider	43
Recommendations	44
Approaches to Defining the Scope of a Guideline	44
Current Practice at CDC	45
Clarifying the Method and Analytic Framework	47
Points to Consider	48
Recommendations	48
Approaches to Method and Analytic Framework	49
Current Practice at CDC	54

Identifying and Synthesizing the Evidence	57
Points to Consider	58
Recommendations	59
Approaches to Synthesizing Evidence	59
Aiding Group Interaction and Decision Making	63
Points to Consider	64
Recommendations	64
Approaches to Aiding Group Interaction	65
Current Practice at CDC	73
Choosing a Group Process Method	73
Identifying a Research Agenda	75
Example 1	76
Example 2	76
Recommendations	77
Updating Guidelines	79
Points to Consider	79
Recommendations	80
Writing the Guideline	81
Points to Consider (Preparing to Write)	82
Recommendations (Author Responsibilities)	82
Basic Components of a Written Guideline	83
Obtaining Critical Reviews and Public Comments	87
Points to Consider	88
Recommendations	88
Approaches to Obtaining Critical Reviews	89
Encouraging the Adoption of Guidelines	91
Issues and Methods	91
Points to Consider	92
References	101
Glossary	107
Abbreviations	113
Appendix A: Desirable Attributes of Clinical Practice Guidelines	115
Appendix B: Instrument for Assessing Clinical Practice Guidelines	117
Appendix C: CDC/ATSDR Federally Chartered Advisory Committees	181
Appendix D: Accessing CDC Prevention Guidelines Database	185

Tables and Figures

Table 1. Model of CDC Guideline Development Process, Participants, and Products	10
Table 2. Rating Scales: From USPSTF	56
Table 3. Characteristics of Group Process Method	74
Table 4. Guideline Cross-Referencing Checklist	86
Figure 1. Analytic Framework for Estrogen Replacement Therapy	55
Figure 2. Analytic Framework “Filled In” with Evidence	77
Figure 3. Framework for Encouraging the Adoption of Guidelines	100

This page intentionally left blank.

Executive Summary

CDC GUIDELINES: Improving the Quality is a guide to improve the development processes, content, and value of CDC-sponsored practice guidelines. The guide presents “points to consider” and “recommendations” which should be given careful attention by CDC staff responsible for developing guidelines. To prepare the guide, a working group of CDC staff synthesized the best advice on guideline development methods available in published journal articles and reports. This summary describes the agency’s history with guideline development, each of 13 primary tasks associated with the guideline development process, and important “points to consider” when developing practice guidelines. Table 1 associates each task in the CDC guideline development process with the participants and task-specific products.

CDC develops a broad range of guidelines which offer advice to clinicians, public health practitioners, managed care organizations, and the public on how to improve the effectiveness and impact of public health interventions. Guidelines developed by CDC over the years generally have been highly regarded within and outside the health care and public health communities, and may frequently be considered the standard of preventive health care. CDC guidelines are currently developed by chartered advisory committees, ad hoc groups, and CDC staff. The processes used vary widely across topics which generally reflect the agency’s programmatic responsibilities. The rationale for recommendations ranges from expert opinion to “best available evidence” as determined by explicit rules of evidence.

The authors recognize that a single approach to guideline development at CDC cannot accommodate all situations because of the tremendous range of topics, varying levels of scientific data, and urgency for development of some guidelines. However, CDC must do better for the following reasons. First, the process for developing the guidelines and the rationale for the recommendations are not always made clear to the user. Second, the format and content of CDC guidelines have never been standardized. Third, plans for disseminating the guidelines are sometimes not formulated in advance. Fourth, the extent to which CDC guidelines are followed varies greatly. Fifth, there is no standardized approach to assessing the impact of CDC guidelines; although many CDC programs have used surveillance systems, surveys, or other indicators to measure impact.

The CDC guideline development process is divided into 13 discrete tasks.

1. Planning and coordinating the process. Define the objectives, mobilize resources, and oversee activities to ensure timely and efficient achievement of defined objectives.

Recommendations:

- Identify and summarize existing guidelines on potential topics.
- Select a guideline topic.
- Select a panel structure with guidance from OGC.
- Select panel members and chairperson.
- Define the purpose and scope of the guideline.
- Collect and synthesize evidence.
- Devise a method to deliberate and then make judgments and recommendations.
- Write, edit, and format the guideline for review.
- Provide for peer review, legal review, and input from private citizens who have an interest in the topic. Document the planned review process, including the criteria for selecting reviewers.
- Prepare completed guideline for dissemination and follow up.

2. Assessing user needs. Find out about the nature, extent, and determinants of current practices regarding the candidate intervention, technology, or health problem of interest in the population of primary guideline users; and gauge the level of common concern and consensus on the potential utility of the proposed guideline.

Recommendations:

- Supplement presumptive indicators of need with empirical evidence of need.
- Seek empirical evidence of need from surveillance data, surveys, focus groups, opinion polls, claims, and other administrative databases.
- Weigh the potential costs and benefits of using empirical versus presumptive indicators of need.
- Consider the cost, feasibility, and potential utility of telephone surveys and focus group techniques for collecting empirical evidence of need.

- Give first priority to meeting the needs of the primary audience for most CDC-sponsored guidelines—one or more categories of public health practitioners.

3. Choosing guideline topics. Decide on the topics that should be given priority for guideline development and explain the rationale for each choice.

Recommendations:

- Adopt a formal, explicit topic selection process.
- Document the procedures and criteria for selecting topics.
- Consider appropriate criteria for topic selection such as controversy, inappropriate current practice, discrepancies between appropriate care defined by meta-analyses and actual practice, practice variation, magnitude of morbidity, mortality, preventability, and cost to society.
- Select topics for which sufficient data are available to provide evidence for a guideline except in unusual circumstances, such as emergencies or for unique new technologies.

4. Selecting Guideline Panels. Choose a panel structure, and participants who will assess the scientific evidence and formulate guideline recommendations.

Recommendations

- Consider whether guidelines should be developed internally or externally. If externally, determine whether the group must be chartered under the Federal Advisory Committee Act. If a non-chartered group is preferred, consult with the Office of the General Counsel for guidance. The type of group used may impact which methods of group interaction and decision-making will be appropriate.
- The selection process for panel members will vary depending on the type of panel being used. Advisory Committee members are selected through a formal process, and are appointed as Special Government Employees or representative members. The selection of regular employees and other experts for non-chartered panels will be accomplished in a less formal process.
- Choose panelists who represent relevant technical disciplines rather than specific organizations.
- Include panelists who serve as liaisons with affected private and governmental groups to assure a voice for the relevant viewpoints, if undue influence can be avoided.

- Ensure that panelists who are experts on a particular technology or practice give appropriate weight to evidence that contradicts a publicly stated point of view.
- Experts in reviewing and weighing evidence, who may or may not be subject matter experts, may be in the best position to consider the evidence presented.
- Outside experts (non-panelists) should be solicited to assure that all relevant scientific evidence was considered and interpreted correctly.
- Where applicable, Individuals must comply with the Government Ethics Laws and Standards of Conduct with regard to conflicting financial interests and other appearances of bias. For non-employees, procedures should be established to identify and appropriately deal with individual conflicts of financial interest or other potential sources of bias.
- Invite panelists with vested interests in a topic to serve as non-voting resources to assure completeness of evidence presented and considered.
- Include panelists with methodologic expertise in assessing scientific evidence, and representatives of likely primary users of the guidelines.
- Include consumers in appropriate situations.
- Document procedures and criteria for selecting panel members.

5. Defining the scope of guidelines. Delimit the target population, outcomes, and interventions which are (1) of greatest interest to specific practitioners, the public, and other users, (2) most amenable to clarification by means of systematic assessment and synthesis of scientific evidence, and (3) capable of being fully explored and resolved into clear and specific advice within the limitations of time and resources.

Recommendations:

- Planning staff should make preliminary decisions about target population, outcomes, interventions, best-practice criteria, potential users, time and resources, and relevance of existing guidelines early in the planning process.
- Planning staff should assess expected quantity of evidence of effectiveness for alternative guidelines whose scope varies from narrower to broader.
- The official decision-making body (panel or task force) should accept or revise the preliminary decisions made by the guideline planning staff as soon as possible after the former is convened.

- Guidelines should address issues of safety and effectiveness of relevant high-volume practices in widespread use.
- Consider implementation issues (cost, cost-effectiveness, staffing, insurance coverage, and patient preferences) whether or not they are used to define best-practices.

6. Clarifying the method and analytic framework. Justify the use of one or more guideline development methods (informal consensus, formal consensus, evidence-based, or explicit) and describe or illustrate the chain of causal reasoning which links the recommended health practice to the desirable health outcomes in a defined individual or population by means of credible evidence of effectiveness.

Recommendations:

- Base prevention guidelines on relative benefits, harms, and costs, whenever possible.
- Consider evidence-based methods (arguably more valid and credible) in preference to consensus methods (informal or formal), if both are feasible.
- Train participants, especially for evidence-based methods, as early in the guideline planning process as is feasible.
- Make the services of a guideline development methodologist available to the panel.
- Prepare and use a written analytic framework (narrative, algorithm, decision model, etc.).
- Use a suitable form of exposition to describe the logical link between practice recommendations, evidence of effectiveness, and desirable health outcomes.
- Consider the rules and regulations of other agencies when developing guidelines in areas of common interest. Specifically, CDC staff and expert panels working on vaccines, drugs, and medical devices, should be trained to use the biostatistical regulatory model of the Food and Drug Administration (FDA).

7. Identifying and synthesizing the evidence. Seek, collect, and assess the quality and quantity of empirical evidence of the effectiveness of a proposed intervention for ensuring a desirable health outcome in a defined population. Empirical evidence of effectiveness must be systematically identified, synthesized, and documented using methods that minimize bias and maximize precision.

Recommendations:

- Develop all CDC guidelines on the basis of data synthesis including a systematic review of the literature, and where relevant and feasible meta-analysis.
- Panelists should be involved in specifying the project format, responsibilities, tasks, and questions to be addressed. Panels should address specific and manageable questions.
- Panelists should be provided with the most comprehensive scientific data possible. A summary description of the available studies should be provided and cited in the final panel statement.
- Document process methodology, facts, assumptions, estimates, criteria for findings, and rationale for recommendations. Include estimates of outcomes expected if the panel's recommendations are followed.

8. Aiding group interaction and decision making. Use formal techniques to maximize the contributions of all participants in the guideline development process. Group interaction methods add clarity and explicitness to group decision making, balance power among participants, minimize bias, and ensure documentation of decision rules and products.

- Provide appropriate training in group interaction techniques for CDC staff and panelists.
- CDC staff should plan for group interaction activities, assigning activities to groups which are appropriate for group work.
- Guideline planning staff should select group interaction techniques appropriate to the specific tasks to be accomplished and the nature of the group or panel.
- Circulate materials in advance for review. (For specific techniques such as Delphi, follow protocol.)
- Use skilled facilitators who are trained in the group process technique selected for the group activity, not invested in a particular outcome, and perceived as neutral and professional by group members. A facilitator with an understanding of the particular topic, but not necessarily an expert, may also serve as the working group coordinator.
- Specify an operational definition of consensus as well as how to present less than full agreement of the panel's findings.
- Maintain good written documentation of the group interaction process; and keep in mind that such documents are in the public domain. Use group interaction methods which produce written documentation.

- Avoid voting as a method of seeking consensus. A better approach is to rank ideas and to encourage participants to react and add to the work of others. If consensus is appropriate, use mediation techniques. If consensus is desired, consider using mediation techniques such as those of negotiated rule making.

9. Identifying a research agenda. Identify a list of study questions or areas of inquiry that should receive high priority for scientific investigation and funding. The primary source of study questions is the analytic framework, its assumptions, and presumed causal linkages for which existing evidence is inadequate.

Recommendations:

- Identify and clarify the relative importance of controversies or gaps in knowledge about desired practices and the best implementation strategies.
- Recommend research needed to resolve issues of importance, especially those for which full agreement was not reached.
- Establish a mechanism for periodically assessing progress in finding answers to key questions on the research agenda.

10. Updating the guideline. Decide on a timetable for revising the guideline to reflect new scientific knowledge. The timetable is determined by the strength of current supporting evidence and expectations about new discoveries, effectiveness of the guideline, and changes in the practice environment.

Recommendations:

- Include a statement, based on the best available information at the time the guideline is issued, that indicates a timetable for revisiting the guideline.
- Review and update the guideline when new evidence suggests that its recommendations are incorrect, ineffective, or can be strengthened.
- Consider the timetable for updating guidelines in relation to the research agenda.

11. Writing the guideline. Prepare a document written with unambiguous language and easy-to-follow logic that provides clear recommendations and documents the rationale on which the recommendations are based.

Recommendations:

- Employ writers and editors who are familiar with the subject matter (medical or scientific writing) or who are able to comprehend, interpret, and explain it after a short learning period.
- Match the document's length, format, layout, and style with the audience needs, subject matter, intention of the guideline, and dissemination vehicle.
- Develop and implement a time-phased work plan to keep the project on schedule.
- CDC public health practice guidelines should address certain basic components: the relevance of the health problem, the magnitude of the problem, the nature of the intervention, the guideline development methods, the strength of the evidence, the cost effectiveness, a discussion of implementation issues, evaluation issues, and the recommendations of others. Provide a brief but complete summary and the name of a contact person.

12. Obtaining critical reviews. Obtain input from content and policy experts, practitioners, advocates, and the public about the scientific accuracy, completeness, and ease of implementation of the draft guidelines. Revising the draft guideline in response to such feedback from potential users and other interested persons can increase the credibility of the guidelines.

Recommendations:

- Use experts on subject matter and methods to review draft guidelines and alert developers to challenges in implementation, credibility, and perceptions of utility.
- Involve a multidisciplinary working group, taskforce, or expert panel which represents all stakeholders to encourage acceptance of the guidelines.
- Seek review by relevant content experts to ensure epidemiological, statistical and clinical validity as well as by relevant organizations and agencies to provide broad input on content and policy issues.
- Arrange for public input, if appropriate.

13. Encouraging adoption of the guideline. Ensure that the guideline is appealing, widely disseminated, and encourages potential users to accept and implement the recommended practices.

Issues and Methods:

- To ensure **relevance**, clarify the users and populations for whom the guidelines apply.

- Describe the health **importance** of the proposed interventions or practices from a range of relevant perspectives.
- To ensure **credibility** and improve **confidence** in the guideline, describe the methods, scientific evidence, participants (individuals and institutions), and relationship to existing recommendations.
- To enhance **feasibility** of adoption, aim to address a single, well-defined practice (when possible), increase user performance at all levels, and allow practitioner judgment that is consistent with the evidence of effectiveness in different populations.
- To highlight the **value** of implementing a guideline, present the costs and benefits from all relevant perspectives (societal, public health clinics, local government, managed care organizations, private practitioners); encourage purchasers and providers of care to establish financial and administrative incentives for use of the guideline; and consider the legal implications.
- Plan and implement active dissemination strategies, when possible, including the use of local information networks, respected colleagues, consumer information campaigns, and administrative mechanisms; keep the electronic Prevention Guidelines Data Base current as a venue for guideline dissemination (see Appendix D).

Table 1. Model of CDC Guideline Development Process, Participants, and Products

Process	Participants	Product(s)
Planning & coordinating	CIO staff	Plan & budget
Assessing user needs	CIO staff; Constituents	Statement of user needs
Choosing a topic	CIO staff; Constituents; Panel chair; Methodologist	Initial topic statement
Selecting a panel	CIO staff; Constituents ;Panel chair;Methodologist	Roster of panelists and alternates
Defining the scope	CIO staff; Panelists; Methodologist	Statement of guideline scope
Clarifying the method & analytic framework	CIO staff; Panelists; Methodologist; Consultant (subject matter expert)	Analytic framework linking proposed actions,desirable outcomes, and evidence of effectiveness.
Identifying & synthesizing the evidence	CIO staff; Panelists; Methodologist; Consultant (librarian)	Summary Evidence Tables
Aiding group interaction and decision making	CIO staff; Panelists; Methodologist Consultant (group process expert)	Decision rules; Minutes of group deliberations. Draft recommendations and rationale
Identifying a research agenda	CIO staff; Panelists; Methodologist Consultant (subject matter expert)	List of important unanswered questions in order of priority. Strategic plan for filling knowledge gaps in order of priority.
Updating the guideline	CIO staff; Constituents; Panel chair; Methodologist	Update plan and schedule.
Writing the guidelines	CIO staff; Panelists; Methodologist; Consultant (writer/editor)	Outline of primary guideline document. Drafts of recommendations and supporting documents.
Obtaining critical reviews & public comment	CIO staff; Panelists; Methodologist; CDC/ADS;CDC/OPPE;CDC/OGC	Final draft of guidelines and supporting documents
Encouraging adoption	CIO staff; Panelists; Methodologist; Consultant (publicist)	Publication and promotional plan.

Introduction

The purpose of this guide is to encourage CDC staff responsible for developing practice guidelines to (1) take a standardized and comprehensive approach to guideline planning and development, (2) document the scientific rationale for the public health actions being recommended, and (3) make both the process for development of the guideline and the rationale for the recommendations clear to the end user. The authors hope that guideline developers will use the document to:

- Encourage ongoing self-assessment of the agency's progress toward increasing excellence in guidelines development.
- Train CDC staff and participants in the guideline development process.
- Establish a common language and frame of reference for planning and implementing the guideline development process.
- Remind guideline developers of important issues to consider so that the resulting guidelines will be optimally effective in influencing public health practice.
- Guide the establishment of research priorities in the area of guidelines development methodology.

The guide was developed in response to the growing recognition that:

- The science and art of guidelines development are advancing rapidly.
- Guidelines are expected to play an increasingly important role in efforts to improve the effectiveness and cost-effectiveness of public health practice.
- The utility and effectiveness of CDC-developed practice advisories are likely to increase if they adhere to a standard of excellence that reflects the state-of-the-art and science of guideline development.

The guide will be disseminated in this and other formats through a variety of communications channels including professional conferences and workshops.

In April, 1995, the Associate Director for Science (ADS), Office of the Director, Centers for Disease Control and Prevention (CDC), and the Chief, Prevention Effectiveness Activity (PEA), Office of the Director, Epidemiology Program Office (OD/EPO), began working with repre-

Purpose and Utility

Mandate and Process

sentatives of the Centers, Institutes, and Offices (CIO) to prepare this guide on guideline development methods.

A CDC-wide Guidelines Methodology Working Group (GMWG) was established (members are listed on page iv) and held its first meeting April 18, 1995. Subsequently, the committee met once every two to four weeks to develop a first draft of the guide. The guide was reviewed (reviewers are listed on page v) and revised based on feedback from CDC leadership, CDC technical staff whose work is expected to be influenced by its content, and selected outside experts on guideline development methods.

Audience and Scope

The primary audience for this document is staff of the Centers for Disease Control and Prevention (CDC) who are responsible for developing public health practice guidelines. However, CDC pursues its mission in partnership with governmental and non-governmental health agencies at the federal, state, and local levels and with health care providers working in a variety of clinical settings, including managed care organizations. Consequently, staff members responsible for implementing public health guidelines in partnership settings are a secondary audience, to ensure that our partners are aware of the principles and frames of reference which shape CDC's approach to the development and sponsorship of practice guidelines.

Other secondary audiences also may include (1) members of advisory groups, convened to develop guidelines, who may need a common framework for pursuing their charge; and (2) groups which develop guidelines that may affect the public's health whether or not the documents are prepared in partnership with CDC.

The scope of this document is limited to practical and scientific issues of particular relevance to the CDC's mission of providing national leadership in developing practice guidelines for population-based health promotion and disease prevention. The practical issues relate to planning, implementing, supporting and documenting the process, and to communicating the product to potential users. The scientific issues relate to the central analytic tasks of guideline development, viz.:

- Reviewing the literature.
- Recording evidence of prevention effectiveness.
- Weighing the evidence of effectiveness.
- Considering costs and effectiveness of intervention programs.
- Explaining the scientific rationale of recommendations that appear prudent based on the evidence of effectiveness, and perhaps, cost-effectiveness.

This document is not intended primarily to serve as a comprehensive manual on how to develop, disseminate, and implement practice guidelines; nor is it intended to prescribe a “one size fits all” approach to developing guidelines for all problems and in all situations.

In this document, the term guideline, which includes *advisory*, *guide*, *guidance*, *guideline*, and *recommendation*, means “advice about the appropriateness of taking some action to prevent or ameliorate the consequences of a public health problem in a defined population.” This document is referred to as a guide and the product of the guideline development process as a public health practice guideline merely as a semantic device to distinguish between the two documents. Other key terms are defined in the glossary of terms beginning on page 107.

Terminology and Usage

This page intentionally left blank.

Background

Members of the GMWG believe that recommendations which are linked to strong scientific evidence of effectiveness by clear and convincing logic are more likely to be implemented and to improve health than recommendations based entirely on current practice, expert opinion, or anecdotal experience. We recognize, however, that evidence of effectiveness may not always exist especially when dealing with new medical devices, emerging infections, or recently discovered diseases. Because the literature reveals little empirical evidence about what makes guidelines useful and effective (Field and Lohr, 1992; p 196), we believe further that:

- Use of this document will lead to improved guideline development capability among CDC staff over time.
- Improved guideline development capability among CDC staff will lead to improved CDC-sponsored public health practice guidelines which display the **eight attributes of good guidelines** (Appendix C)(Field and Lohr, 1990; Field and Lohr, 1992).
- Improved practice guidelines will enable individuals and organizations (health departments, managed care organizations, and community-based organizations) to develop and implement more cost-effective prevention programs.
- More cost-effective prevention programs will lead to improved community health status at affordable levels of resource use.

The sequence of activities which leads to a set of guidelines includes activational work (organizing a group to develop practice guidelines), analytic work (reviewing scientific evidence, developing recommendations, and explaining the rationale), editorial work (writing the document) and diffusion work (the review, dissemination, and adoption of guidelines by the user audience). Woolf and others contend that the analytic work is clearly the most important and complex task affecting the quality of practice guidelines (Woolf, 1994).

Some background information about the agency is important to understand guidelines development at the CDC and the Agency for Toxic Substances and Disease Registry (ATSDR). CDC and ATSDR are agencies of the US Public Health Service within the Department of Health and Human Services (DHHS). These sister agencies have a common leadership, complementary missions, and a single organizing vision. For CDC that vision is “Healthy people in a healthy world through prevention.” In 1994, CDC’s name was changed to “Centers

Assumptions

Public Health Practice at CDC

for Disease Control and Prevention” to underscore its focus on preventive interventions. CDC is primarily a public health agency, i.e., it considers the community as the patient, identifies important community health problems, and seeks to develop community interventions to address those problems. Some of CDC’s areas of interest are (1) infectious diseases, (2) chronic diseases, (3) environmental health, (4) occupational health and safety, (5) injury prevention and control, (6) health statistics, and (7) international health.

In addition, CDC trains public health workers in epidemiology and other applied public health areas and responds to public health emergencies, e.g., epidemic investigations, Health Hazard Evaluations, and Crisis Response Teams. CDC accomplishes its mission through a variety of activities including research, public health surveillance, outbreak investigation, evaluation of interventions, public health capacity building, and national leadership. National leadership includes developing guidelines. Approximately 80% of CDC’s budget is spent extramurally through grants, cooperative agreements, and contracts. Congress targets much of CDC’s money to program-specific areas (e.g., immunizations, human immunodeficiency virus (HIV), sexually transmitted diseases (STD), tuberculosis (TB), and lead poisoning, etc.) and to specific entities such as health departments and community-based organizations.

CDC does not operate independently; it places strong emphasis on partnerships. These partners include (1) state and local health, education, and public safety departments, (2) other federal and international agencies, (3) academic institutions, (4) labor and management, (5) health maintenance organizations (HMO), hospitals, and other health care provider groups and institutions, (6) philanthropic foundations, (7) professional societies, (8) voluntary and community-based organizations, (9) physicians and other health care providers, (10) laboratory personnel, (11) and the general public. The list is neither exhaustive nor mutually exclusive.

Public Health Practice Guidelines

The IOM defines clinical practice guidelines as “systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances.”

Because public health practice is not limited to clinical circumstances, the IOM definition is expanded here to include public health circumstances. For the purposes of this guide, public health practice guidelines are defined as systematically developed statements to help policy-makers, public health practitioners, clinical practitioners, health agencies, and the public decide the appropriate actions to (1) promote health, (2) prevent disease, disability, and premature death, and (3) improve quality of life of members of a defined population. Appropriate actions may target individuals or whole communities at risk for adverse health outcomes. Public health practice guidelines may refer to

clinical preventive services which are defined as recommendations for clinical practice on preventive interventions—screening tests, counseling interventions, immunizations, and chemoprophylactic regimens—for target conditions among asymptomatic individuals of all age groups and risk categories (U.S. Preventive Services Taskforce, 1989).

The IOM proposes that “clinical practice guidelines have real potential to help clarify the knowledge base for clinical practice and to improve the quality and effectiveness of medical care” (Field and Lohr, 1992). A review of the most rigorous studies of the effectiveness of strategies to improve the dissemination and application of health care knowledge by physicians concludes that strategies then in common use were incompletely tested (Lomas, 1988). Although empirical evidence of effectiveness similarly is lacking, it is reasonable to expect that public health practice guidelines have potential to improve the quality and effectiveness of public health practice.

The IOM asserts that clinical practice guidelines which possess eight desirable attributes are more likely to be perceived as trustworthy, usable, and effective in achieving desired health outcomes. Good practice guidelines are demonstrably (Field and Lohr, 1992)(Appendix A):

- **Valid**—supported by strong evidence linking recommendations to outcomes.
- **Reliable or reproducible**—prepared using procedures and decision rules that would lead different experts to the same conclusions based on the same evidence.
- **Applicable**—useful in populations that potential users would consider relevant to their practices.
- **Flexible**—allowing for practitioner judgment and patient preferences.
- **Clear**—presented in unambiguous language and easy-to-follow logic.
- **Multidisciplinary**—prepared with input from relevant disciplines and stakeholders.
- **Up to date**—reflecting the most recent evidence.
- **Documented**—published along with explicit statements on assumptions, process, rationale, evidence, and decision rules.

Rationale for Using Guidelines in Public Health Practice

Desirable Attributes of Guidelines

Guidelines Development at CDC

CDC develops guidelines to influence others to take actions which will reduce preventable morbidity and mortality from diseases and other health conditions. The specific “triggers” for developing a specific guideline are numerous and variable. Frequently, a guideline is developed in response to an emerging health problem such as hantavirus pulmonary syndrome or multidrug-resistant tuberculosis. Guidelines also are developed in response to requests from constituents such as health care providers or public health officials. The passage of a new law or the development of regulations might prompt the development of a guideline. Another trigger might be new surveillance data; several vaccine recommendations were revised in response to changes in disease epidemiology identified by surveillance data. The results of a research study or series of studies might precipitate the development of a guideline, for example, the appropriate use of folic acid for prevention of neural tube defects. Some guidelines were developed because of a new technology, e.g., varicella vaccine. This list is not exhaustive nor are the triggers mutually exclusive.

Broadly speaking, the audience for a guideline is anyone who can take the recommended actions. This audience includes many groups and individuals listed on page 16 (see section on public health practice at CDC).

Different approaches are used to develop guidelines depending upon the circumstances. Some guidelines are developed by Federally chartered advisory committees, e.g., the Advisory Committee for Immunization Practices (ACIP), the Advisory Committee for the Elimination of Tuberculosis (ACET), and the Hospital Infection Control Practices Advisory Committee (HICPAC). Some of these committees are composed of individuals appointed by the Secretary, DHHS, others include ex-officio members from other government agencies, and liaison members from non-federal organizations with an interest and expertise in the topic. Some guidelines are developed by ad hoc advisory groups. Still others may be developed by CDC staff.

Many of CDC guidelines are developed in collaboration with other organizations, especially other Public Health Service agencies such as the National Institutes of Health, the Food and Drug Administration, and the Agency for Health Care Policy Research. Some guidelines are published in draft form in the Federal Register for comment. At times, public meetings are held at CDC to obtain comments from the public on a proposed guideline.

The guidelines development process varies from topic to topic, program to program, and committee to committee. Generally speaking, CDC staff conduct a literature and data review and summarize it for the benefit of outside advisors and other CDC staff. Meta-analyses, or systematic reviews, are sometimes performed. Based on these summaries, policy options, treatment options, etc., are debated. Decision analyses, cost-effectiveness analyses, or other quantitative approaches

to decision making are being used increasingly at CDC. Other inputs include legal review, public comments, societal values, and expert judgments about feasibility.

Characteristics common to virtually all CDC guidelines include an openness to input and extensive and repeated review by people both within and outside of CDC. Most CDC guidelines are published in the Morbidity and Mortality Weekly Report (MMWR). The topics of CDC guidelines reflect the agency's program-specific funding priorities which include (1) infectious diseases such as HIV infection, vaccine preventable diseases, sexually transmitted diseases, and tuberculosis and risk factors for antimicrobial resistant infections; (2) chronic diseases such as breast and cervical cancer and diabetes and risk factors for chronic disease such as poor nutrition, smoking, and lack of exercise; (3) injury prevention and control; (4) reproductive health, maternal and infant health; (5) environmental health; (6) occupational health; (7) international health; and (8) epidemiology and health statistics.

CDC develops a broad range of documents that could be called guidelines. Guidelines developed by CDC over the years generally have been highly regarded within and outside the health care and public health communities, and may frequently be considered the standard of preventive health care. However, the format and content of CDC guidelines have never been standardized. Plans for disseminating the guidelines are sometimes not formulated in advance. The extent to which CDC guidelines are followed varies greatly. There is no standardized approach to assessing the impact of CDC guidelines; although many CDC programs have used surveillance systems, surveys, or other indicators to measure impact.

CDC guidelines are currently developed by chartered advisory committees, ad hoc groups, and CDC staff. The processes used vary widely across topics which generally reflect the agency's programmatic responsibilities. The rationale for recommendations ranges from expert opinion to "best available evidence" as determined by explicit rules of evidence. The process for developing the guidelines and the rationale for the recommendations are not always made clear to the user. For these reasons, CDC must do better.

This guide is offered as an aid to CDC staff for improving the guidelines planning, development, and dissemination process, recognizing that with the tremendous range of topics, varying levels of scientific data, and urgency for development of some guidelines, a single approach to guideline development at CDC cannot accommodate all situations.

Improving the Quality of CDC Guidelines

Ethical Issues

Specific ethical issues may arise in the context of any of the 13 guideline development tasks. For example, the issue of conflict of interest is addressed in the sections of this document dealing with selecting guideline panels (page 39) and aiding group interaction (page 65). Moreover, the professional conduct of CDC staff engaged in guideline development are guided by standards of ethical conduct for employees of the executive branch of the federal government (Title 5, Part 2635, Code of Federal Regulations).

In general, however, CDC staff involved in guideline development must be guided by the basic moral principles of (1) respect for autonomy in decision-making of patients, clients, and practitioners, (2) beneficence—enhancing the welfare of others, (3) nonmaleficence—avoiding harm to others, and (4) justice—equity in the distribution of benefits and risk (Hahn, 1994). Weed argues that advocacy in the form of public health recommendations can be justified in terms of the principle of beneficence found in the guidelines (Weed, 1994). With respect to nonmaleficence, clinical practitioners, patients, payers, and others have continued to voice concerns about the potential harms of clinical practice guidelines. The catalogue of potential harms include (1) insensitivity to differences in patient populations and practice settings, (2) threats to professional autonomy, income, credentials, and hospital privileges, (3) disincentives to independent thinking and innovation, and (4) inappropriate rationing of health care. (Woolf, 1993). In the mid-1960s, Hill argued that judgment is important in making public health recommendations; and in his mind, different levels of evidence were required for public health action, depending on who was affected by those actions (Coughlin and Beauchamp, 1996).

Controversy and Consensus

There are numerous controversies surrounding guidelines development. For example, some contend that the terms guide, guideline, guidance, and recommendation have different shades of meaning in common usage. Others argue that there is no substantive difference in the meaning of these words.

Controversy also surrounds the issue of distinguishing clinical guidelines from public health practice guidelines. Some argue that clinical guidelines, which relate to the diagnosis and treatment of illness in individuals, are qualitatively different from public health practice guidelines, which relate to prevention and health promotion among populations. Others contend that clinical guidelines are a subset of public health practice guidelines, which relate to all phases of health promotion and disease prevention for individuals and populations.

In parallel with these ongoing controversies about language and context, a consensus is emerging among CDC staff and external partners that future guidelines developed at CDC can be improved while maintaining a balance between the complementary goals of scientific rigor, feasibility in terms of time and development costs, and practicality in use.

Planning and Coordinating the Process

The way guidelines are developed can strongly affect their potential for use by the intended audience. Thus, the entire process from introductory decisions through revisions and implementation requires careful planning and coordination (Field and Lohr, 1992). Certain steps are central to guideline development. Groups, however, may differ somewhat on the emphasis that the steps are given and the sequence in which they are performed. The steps in planning and coordinating guideline development can be categorized broadly as introductory decisions (about the topic, the authors, and its purpose), assessment of the public health or clinical appropriateness of the proposed guideline (based on evidence, expert opinion, and decision rules), assessment of public policy issues (related to resource limitations and feasibility), and guideline document development and evaluation (Woolf, 1992).

Answers to the following questions can help plan and coordinate the process:

- Is the topic appropriate for developing a guideline?
- Is there a source of evidence on which to base the guideline?
- Is a technical workplan being developed?
- Is an administrative workplan being developed?
- Have existing guidelines on potential topics been identified and summarized?

The recommendations outlined below provide general suggestions for group planning and coordinating of the process of guideline development.

- Identify and summarize existing guidelines on potential topics.
- Select a guideline topic.
- Select a panel structure with guidance from OGC.
- Select panel members and chairperson.
- Define the purpose and scope of the guideline.
- Collect and synthesize evidence.

Points to Consider

Recommendations

- Devise a method to deliberate and then make judgments and recommendations.
- Write, edit, and format the guideline for review.
- Provide for peer review, legal review, and input from private citizens who have an interest in the topic. Document the planned review process, including the criteria for selecting reviewers.
- Prepare completed guideline for dissemination and follow up.

Approaches to Planning and Producing Guidelines

Historically, CDC-sponsored guidelines have been developed by (1) Federally chartered advisory committees, (2) ad hoc expert panels which are convened to develop a particular guideline, (3) staff of the sponsoring CIO, and (4) some combination of the categories described in items 1-3. Although there are more similarities than differences, the planning and coordinating for guideline development efforts at the CDC vary depending on the development group.

The Federal Advisory Committee Act (FACA) must be kept in mind when considering the creation of a guideline panel which includes non-government employees. **This includes groups convened to develop guidelines or to provide input to the guideline developers.** Generally, the FACA governs the formation and functioning of groups which include non-government employees, convened for the purpose of obtaining consensus advice or recommendations on issues or policies which are within the scope of an agency's responsibilities. Therefore, whether the group is a formally chartered advisory committee or not, may impact which methods of group interaction and decision-making will be appropriate. Chartered advisory committees are free to reach consensus on advice and recommendations to the agency. Other types of groups may not **provide consensus advice and recommendations to the agency. Advice and recommendations may be obtained through the individual input of outside experts.** However, there are group models where consensus recommendations, such as guidelines, may be generated for the public health community, or other external constituencies. If groups are formed outside the FACA, and include non-government employees, it will be necessary to consult with the Office of the General Counsel.

Examples of the steps used by the groups listed in items 1-3 above are outlined below.

Federally Chartered Advisory Committees

ACIP is an example of a federally chartered advisory committee. Guideline topics are chosen by the committee, with assistance from CDC staff. Developments at the Food and Drug Administration (FDA) (e.g., a manufacturer's submission of a product for regulatory approval) may prompt a review of a new vaccine. A memo is circulated periodically to committee members, organizational liaisons, and CDC-wide to elicit topics for consideration. The starting point for ACIP rec-

ommendations may be labeling and package inserts of licensed vaccines or a previous ACIP recommendation on the same subject. That step is followed by:

- A thorough review of the scientific literature (both published and sometimes unpublished) on the immunizing agent and key prevention issues. A systematic review is done usually or a meta-analysis is done sometimes.
- Ascertaining the relevance and quality of published and unpublished data.
- An assessment of the morbidity and mortality of the disease, the safety and efficacy of the vaccine, and the feasibility of its use. The scheduling of immunizations is also considered. Cost-effectiveness and cost-benefit analyses are frequently performed to aid committee decision making.
- An extensive review of existing recommendations by CDC staff, ACIP members, and outside expert consultants. Working groups are often formed to review the research data prior to presentation to the full committee.
- Public comments are solicited during the committee meetings and are considered in the decision-making process.
- All recommendations are subject to an exhaustive review and are made only after extensive dialogue among the committee, liaison members, and other concerned parties. All recommendations are also discussed at public meetings. Some recommendations are published jointly with other recommending bodies, such as the American Academy of Pediatrics, American College of Physicians, etc.

Ad hoc panels are convened to develop guidelines on particular topics.

Ad Hoc Panels

Example: The U.S. Public Health Service (USPHS)/Infectious Diseases Society of America (IDSA) Prevention of Opportunistic Infections Working Group. The USPHS—primarily through the efforts of the CDC and the National Institutes of Health (NIH)—and the IDSA in 1994 recognized the importance of preventing opportunistic infections and the need to consolidate information for health-care providers. In response, these organizations initiated an effort to develop comprehensive recommendations for the prevention of opportunistic infections in HIV-infected persons. The goal was to disseminate information on pathogens that can cause disease in patients with HIV infection and the chemoprophylactic regimens available for preventing disease. Such information about preventing exposure and preventing disease may have been published in journals that are not regularly reviewed by healthcare providers; and some of it may not have been published. Draft recommendations were reviewed by staff from CDC, NIH, and IDSA, as well as by members of other

government agencies, community organizations, physicians caring for HIV-infected persons, and HIV-infected persons themselves.

These recommendations were discussed at a two-day meeting convened by the CDC, NIH, and IDSA in Atlanta in September 1994. Comments were solicited from the public and final recommendations were approved by the USPHS and IDSA. The recommendations were endorsed by the American Academy of Pediatrics, the Infectious Diseases Society of Obstetrics and Gynecology, and the Society of Healthcare Epidemiologists of America.

CDC Staff

CDC staff guidelines are developed in house in response to a particular need and may be developed in collaboration with other groups.

Example: The Guidelines for School Health Programs to Prevent Tobacco Use and Addiction. These guidelines represent collaboration with experts from 29 national, federal, and voluntary agencies along with other leading authorities in the field of tobacco-use prevention. The steps involved:

- An in-depth review of research, theory, and current practice in the area of school-based tobacco-use prevention.
- CDC staff convening meetings of experts from the fields of tobacco-use prevention and education.
- A review of published research.
- Consideration of the conclusions of the National Cancer Institute Expert Advisory Panel on School-Based Smoking Prevention Programs and the findings of the 1994 Surgeon General's Report, "Preventing Tobacco Use Among Young People."

Administrative Issues

In addition to this global planning which must take place, there will also be an assortment of managerial, or administrative, issues which must be decided and handled. While they may also vary to some degree, depending on the group, many of the issues are universal with guideline development. Unless the group is already well organized, like ACIP and HICPAC, certain mechanisms must be brought into place. For example, what will be the source of funding for this project? Is the administration going to be coordinated internally or is a contractor going to be hired to handle it? Will training be required and if so, how will it be handled? Will external consultants be needed during the process? An assortment of other issues also need to be resolved. Provisions must be made for dissemination, review, and clearance—e-mail, fax, conference calls. Publication requirements must be checked and projected costs obtained.

No matter who handles them, the administrative duties will also include the need to arrange meetings and provide for any special requirement (such as slides and audiovisual equipment) that presenters may

need at the meetings. Notebooks will need to be assembled and materials distributed. A tracking system must be put into place. Staff must be identified to develop the agenda for the meetings, keep the minutes, and provide for public and financial disclosure, if needed. Arrangements for press releases and conferences may have to be made sometimes. These issues represent just some of the basic housekeeping duties that may be required within the broader basic steps of guideline development. It is essential that they be considered early on to ensure that the process runs smoothly.

This page intentionally left blank.

Assessing User Needs

The team planning a guideline development effort must justify committing the required resources to the effort. Evidence of perceived pressing needs (for knowledge, skills, and a good reason to change their practices) among potential users of the proposed guideline often is sufficient justification. Assessing user needs means finding out about the nature, extent, and determinants of current practices regarding the candidate intervention, technology, or health problem of interest in the population of primary guideline users. Documenting the perceived educational and other needs of the target audience is a prerequisite for choosing a guideline topic. (See next section.) Also it gauges the level of common concern and consensus on the potential utility of the proposed guideline.

The potential utility of a guideline will vary among categories of users. Public health practitioners, clinicians, patients, administrators, payers, politicians, and other users of practice guidelines have different needs and expectations of what good will result from their use. The use of guidelines often are expected to improve population-based health status, individual health outcomes, and to reduce costs of care, practice variations, public expenditures, and inappropriate care. In the context of this guide, guidelines are perceived as the starting point of planned interventions designed to influence knowledge, attitudes, skills, and professional and organizational behaviors (Woolf, 1993).

Green and Kreuter have identified three categories of factors that influence behavior or professional practice: predisposing, enabling, and reinforcing factors. Predisposing factors include the professional's values, beliefs, attitudes, and perceptions about guidelines and their potential usefulness. These factors account for the professional's motivation to use guidelines and confidence in being able to implement their recommendations. Enabling factors are the necessary skills and resources (finances, staff, space, and educational materials) which the professional must possess in order to successfully implement the guideline's recommendations. Reinforcing factors are the rewards or incentives that are anticipated or that actually follow as a consequence of a particular behavior. They include reimbursements actually received, visible improvements in patient or population health outcomes, support from colleagues, and feedback from patients or clients (Green and Kreuter, 1991, pp 408-416).

Practitioners are unlikely to use guidelines that do not meet a perceived need for credible information. Here we distinguish between a perceived need (the understanding among users that the guideline will be valued) and a socially judged need (the declaration by public health

experts that the guideline should be valued). Indicators of perceived need for a guideline may be presumptive or empirical. Presumptive indicators of need are conventional “triggers” of guideline development efforts at CDC. These triggers were described previously in the background section of this document (See page 18) . Empirical indicators of practitioner needs for new knowledge, skills, and incentives would come from the results of surveys, focus group interviews, opinion polls, and analyses of practice patterns based on claims or other administrative data sources.

Points to Consider

CDC staff may characterize or justify the need for a particular guideline by providing explicit answers to the following questions about the proposed guidelines.

- What is the “trigger” that serves as the presumptive indicator of need for the guideline (an emerging health problem, requests from constituents, passage of a new law or the development of regulations, new surveillance data, results of a research study or series of studies, licensure of a new technology or vaccine, unexplained or inappropriate practice variations, and recognition of an outdated guideline)? Is the guideline required by a legislative, executive, or judicial mandate? Is the guideline needed to provide standards of procedure and quality control in a CDC-funded grant program or other activity (e.g., STD, TB, immunization, prevention of disease during international travel)?
- Are empirical indicators of practitioner needs available (e.g., results of surveys, focus groups, opinion polls, or analyses of practice patterns based on claims or other administrative data sources)?
- Are data available on the nature, extent, and determinants of current practices regarding the candidate intervention, technology, or health problem of interest among the population of potential primary guideline users?
- Is there an existing guideline on the same topic? Is it adequate? How well is it being followed? How will the proposed guideline be different?

Recommendations

These recommendations are based on the assumption that the intent of the proposed guideline is to inform and possibly influence the professional practices of the target users.

- CDC staff should strive to supplement legislative, programmatic, and presumptive indicators of need for guideline development with empirical evidence of practitioner need. Empirical evidence of practitioner need for the proposed guideline should be sought from primary or secondary sources of data derived from surveillance, surveys, focus groups, opinion polls, and analyses of practice

patterns captured on claims or other administrative data bases. Guidelines which address empirically documented practitioner needs are more likely to be used than those which address presumed needs.

- The potential costs and benefits of postponing a guideline development effort pending availability of empirical evidence of practitioner need should be weighed against the costs and benefits (including potential impact on disease prevention) of proceeding on the basis of presumptive indicators of need.
- The cost, feasibility, and potential utility of telephone surveys and focus group techniques for collecting empirical evidence of practitioner needs should be considered.
- Because it is impossible to satisfy all needs and expectations of all potential guideline users equally well, first priority should be given to meeting the needs of the primary audience for a particular guideline. Usually, the primary audience for most CDC-sponsored guidelines is one or more categories of public health practitioners.

Empirical information about the needs of practitioners is relevant to the assurance of at least three of the eight desirable attributes of guidelines. Those desirable attributes are (1) applicability—i.e., covering priority populations and a variety of practice settings; (2) flexibility—i.e., allowing provider judgment and patient preferences; and (3) multi-disciplinary content—i.e., shaped by input from diverse scientific disciplines and affected stakeholders. Assessing the needs of potential users of guidelines is similar to conducting a health needs assessment in a community. The purpose is to determine potential guideline users' perceptions of their own needs, aspirations for the common good, and the role they themselves might play in changing their own professional practices (Green and Kreuter, 1991 pp 44-87).

Both qualitative and quantitative methods can be used to assess user needs.

In most needs assessment situations, applications of the two types of approaches are combined. For example, qualitative methods often are used to improve the quality, validity, and interpretation of data gathered by quantitative approaches (Steckler, 1992).

Qualitative methods rely on techniques adopted from the social sciences to elicit an “insider’s view” (study participants or participant observers) of how members of a group under study perceive their own needs for guidelines, how they are likely to react to such guidelines, and what positive or negative consequences they expect to result from the use of those guidelines. Such methods include community forums or workshops, focus group interviews, nominal group process, key informant interviews, and archival research. The section of this guide en-

Approaches to Assessing User Needs

titled Aiding Group Interaction and Decision Making (See page 67) discusses the strengths and weaknesses of several qualitative methods in the context of eliciting and combining group judgments. Many of the same principles apply when those methods are used to elicit and to characterize the needs of potential users for proposed guidelines.

Community forums or workshops assemble interest groups for intensive meetings to discuss issues surrounding proposed guidelines and to reach an understanding among the participants about needs, scope, format, timetables, and roles. Focus group interviews are informal sessions in which eight to 12 potential users of a proposed guideline are asked to discuss their thoughts and feelings about the issue. Those thoughts and feelings are then used to help clarify the content, delivery, and appeal of the proposed guidelines. The nominal group process or technique elicits written responses to a single question without verbal interaction among group members. (See page 67). The technique would include small groups of five to nine potential guideline users to assess target group perceptions of need and obstacles to meeting that need. This technique reduces the tendency for the more socially powerful to dominate the discussions and to bias the consensus that emerges.

Key informant interviews rely on in-depth, semi-structured interviews with individuals who are selected to participate because they have special knowledge or insight not available to others in their reference group—hence the label “key informant.” Recorded or written transcripts of the interviews are then analyzed to uncover themes and the relative importance of frequently mentioned issues relating to guideline development and use (Hugentobler, 1992). Archival research examines reports, newspaper clippings, correspondence, books, and other documents prepared by persons other than the researcher and kept in archival depositories such as libraries. Archival sources of information about the needs of guideline users are important because they reveal changes in perceptions and practices over time (Wolcott, 1992).

Quantitative methods rely on numerical measurements and statistical techniques to estimate aggregate characteristics of the group under study and to draw inferences which can be generalized to a larger reference population. Quantitative methods include sample surveys and the construction of synthetic estimates from administrative data sources. In the remainder of this section, we describe two examples of the use of surveys to assess user needs.

Examples of Quantitative Methods

In 1995 the Canadian Community Health Practice Guidelines (CHPG) Working Group developed practice recommendations for three public health activities: restaurant inspections, STD partner notification, and immunization delivery. In each program area, a Canada-wide survey of public health units was done to document current practice and variability in practice across jurisdictions. For example, the 1991 practice survey of immunization delivery methods interviewed provincial epidemiologists and other key persons involved in

immunization programs to document province-specific variations in public/private health administration, legislation, monitoring system/coverage rates/surveillance, vaccine management and costs. The results of the survey helped to define the educational needs of public health units regarding immunization service delivery. For example, the survey showed that programs were not being formally evaluated; instead, program managers obtained information from local experience, trial of new practices, and organized research projects (Gyorkos, 1995).

In 1991 the North Carolina Department of Health, Environment, and Natural Resources (DHENR) and the National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP) conducted a telephone survey of counseling and referral practices among a statewide probability sample of primary-care physicians. Respondents (n=514; 58.6% response rate) provided population-based estimates of the proportion of primary-care physicians who counsel and/or refer for treatment patients who smoke, abuse drugs or alcohol, or have diet- or nutrition-related problems. Although the survey was not conducted in support of a specific guideline development effort, the DHENR, the North Carolina Medical Society, and NCCDPHP planned to use the survey results to identify and help address, perhaps through the medium of practice guidelines, the educational needs of primary-care practitioners with regard to health education and preventive services (CDC/MMWR, 1992).

CDC programs rely heavily on presumptive evidence of need for a guideline—i.e., traditional “triggers” as described earlier in this section; empirical indicators of perceived practitioner need for new knowledge and skills rarely are collected routinely or systematically.

Current Practice at CDC

This page intentionally left blank.

Choosing Guideline Topics

Guideline topics should be selected based on the need for guidelines and, secondarily, on the amount and quality of scientific evidence available to make writing guidelines feasible. The optimal guideline topic addresses the impact of a specific technology on a health outcome using the societal perspective. The topic can focus on any number of different issues including the efficacy, effectiveness or cost-effectiveness of a technology, individual service, or population-based service. Recommendations should be based on the evidence linking the intervention with the health outcomes.

Topics may be categorized using any of several categorization schemes and sub-schemes. Selection of a scheme is often obvious as the purpose of the group developing the guidelines has a readily apparent mandate, e.g., ACIP for immunizations or the American Lung Association for tuberculosis. Among possible schemes are:

- **Disease or injury (outcomes)** — chronic disease, maternal and child health, occupational health and safety, and injuries.
- **Risk factor**—behavioral or physiological risk factors (smoking, hypertension, etc.).
- **Type of technology** — immunizations, education programs, toxin control, reminder systems, outreach.
- **Users** — family practice physicians, nurses, educators, hospitals, managed care organizations
- **Target population** — infants, inner city residents, elderly.
- **Some combination of the above**

The IOM has recommended that an explicit process be followed in setting priorities for guideline development (Field, 1995). The actual topic selection process can be accomplished by one or more of the following:

- **A legislative, executive, or judicial mandate:** For example, Congress may require the development of guidelines directly or based on some criteria, such as Medicare resources expended.
- **Agency priorities and responsibilities:** CDC programs may need guidance for their own operations or to assure quality of services. STD diagnosis and treatment guidelines, for example, are regularly issued to improve the management of those diseases. Other agency responsibilities, such as immunization, international traveler's

health, hospital infections, and school health education, lead directly to guideline development.

- **Process involving a task force or advisory committee:** Certain advisory groups are constituted explicitly to prepare guidelines. The Advisory Committee on Immunization Practices (ACIP) is an example. Advisory groups may establish their own topic selection process.
- **Surveys:** Formal solicitation of topics from users, experts, employers, or managed care organizations.

Points to Consider

The need for systematic thinking about criteria for topics requires panels to have clarity of purpose. Prioritization of topics can be done on the basis of qualitative as well as quantitative criteria. Quantitative criteria to consider include:

- **Morbidity, mortality, or quality of life:** The public health impact of the disease or injury under consideration may be assessed by incidence, prevalence, mortality, severity, disability, and other morbidity.
- **Preventability, treatability, or curability:** Is there a technology which can improve the health outcome?
- **Economic impact:** Is this a costly outcome based on medical care costs and losses of productivity?
- **Cost of intervention:** What is the cost of the prevention strategy or technology in the aggregate or on a unit cost basis?
- **Controversy:** There are often controversies about effectiveness and cost effectiveness of different approaches based on discipline (behavioral, clinical, epidemiological, or environmental emphasis) or schools of thought. Guidelines may provide the evidence to reduce disparities and increase effectiveness.
- **Availability of evidence:** The number and quality of studies which have been performed for different technologies vary widely from extensive randomized trials to purely descriptive information. Evidence-based guidelines require the availability of high quality studies, but not necessarily randomized controlled clinical trials.
- **Variation in practice:** When clinical practice varies from community to community there is also often variation in costs and health outcomes. Such variation in practice, costs, and outcomes may be due to high levels of inappropriate care. In these circumstances, guidelines may lead to more widespread use of the more effective or cost-effective technique. While variations have not been examined as closely for public health practices, they may also suggest technologies for which guidelines may improve outcomes.

- New versus established interventions: In general, guidelines are more useful for newer technologies which have not become part of standard practice. Sometimes, however, guidelines on making better use of old technologies are equally useful.
- A formal, explicit topic selection process should be adopted, when possible.
- The procedures and criteria for selecting topics should be documented.
- Controversy, inappropriate current practice, discrepancies between expected practice (defined by meta-analyses) and actual practice, and unexplained practice variation are appropriate criteria for topic selection.
- The magnitude of morbidity, mortality, preventability, decreased quality of life, and cost to society are appropriate criteria for topic selection.
- Topics should not be selected unless sufficient data are available to provide evidence for a guideline except in unusual circumstances, such as rapidly emerging health problems or the development and imminent use of unique, new, potentially life-saving technologies.

Recommendations

This page intentionally left blank.

Selecting Guideline Panels

CDC's Office of the General Counsel (OGC) should be consulted prior to the formation of a panel or workgroup to develop guidelines. Consider whether guidelines should be developed internally or externally. If externally, determine whether the group must be chartered under the Federal Advisory Committee Act. If a non-chartered group is preferred, consult with the Office of the General Counsel for guidance. The type of group used may impact which methods of group interaction and decision-making will be appropriate.

The selection process for panel members will vary depending on the type of panel being used. Advisory Committee members are selected through a formal process, and are appointed as Special Government Employees or representative members. The selection of regular employees and other experts for non-chartered panels will be accomplished in a less formal process.

The composition of guideline panels can shape the recommendations themselves. From the size of the panel to the characteristics of the members, each decision may affect the group dynamics or potential bias of the group as a whole. The importance of an unbiased panel increases as the strength of the scientific evidence declines (US Congress, Office of Technology Assessment (OTA), 1994, p.11).

One of the primary goals in selecting panel members is to minimize threats to the validity of the conclusion. Bias occurs easily, often as a result of professional relationships to the technology under consideration. Organizational or professional representatives may well become advocates for specific organizational perspectives. It is only natural that a surgeon would be inclined to look more favorably on indications for surgery, or that someone managing a screening program would view the evidence for screening more favorably. There is evidence that surgeons find more carotid endarterectomy cases appropriate than does a multidisciplinary team (Leape, 1992). Although such bias has not been documented for population-based recommendations, public health professionals can be anticipated to have preconceived preferences which might bias the interpretation of information.

- Possible sources for panel members include:
 - Well known experts in relevant methodologies (decision analysis, meta-analysis, cost-effectiveness) the particular discipline, or subject matter.

Points to Consider

- Representatives of specific organizations (e.g., American Association of Health Plans, American Academy of Pediatrics, American Heart Association, American Medical Association, schools of medicine and public health, and universities)
 - Researchers in relevant scientific fields (from universities, academic medical and public health centers)
 - Representatives of one or more professional groups (e.g., physicians, nurses, podiatrists, educators)
 - Public health officials (e.g., local, state, CDC, PHS)
 - Consumers
 - Other interest groups (e.g., groups representing women, racial and ethnic minorities, the elderly, the disabled, and others who may be disproportionately affected by the health problem or proposed intervention)
- Will the proposed panel members be able to (1) objectively review and assess the quality of scientific evidence, (2) have familiarity with the prevention strategies under consideration, and (3) participate constructively in group processes?
 - Will those being considered result in a multidisciplinary group?
 - Do those being considered represent the range of culturally diverse populations which will likely be users or beneficiaries of the guidelines?
 - Does any individual being considered have a stated opinion (published articles, editorials, or speeches) on the topic? If so, does the entire deliberative body represent a range of opinions? And is each out-spoken individual also open-minded and willing to listen to other opinions?
 - Those who are most knowledgeable about the technology being assessed are usually the individuals who have extensive experience, publications, and well-established points of view. While they may be open to considering evidence, they may hold evidence to a different standard if it contradicts a publicly stated point of view.
 - Those who are experts in reviewing and weighing evidence but are not subject matter experts may be in the best position to consider the quality of evidence presented.

Recommendations

In selecting panel members, there are tensions between the need for objectivity in reviewing evidence and the need for the input of experts with extensive experience and established positions on the guideline topic. The following recommendations should be considered when constituting a panel:

- Consult OGC prior to the formation of a panel or workgroup to be sure that the proposed process or deliberations do not violate the Federal Advisory Committee Act.
- In general, panelists should represent relevant technical disciplines regardless of whether they represent specific organizations. Panelists who serve as non-voting liaisons with affected private and governmental groups often are needed to assure a voice for the relevant viewpoints, if undue influence can be avoided.
- Regardless of who is on a panel, outside technical experts should be solicited to assure that all relevant scientific evidence was considered and interpreted objectively.
- Panelists should be asked to disclose and discuss possible conflicts of interest and biases.
- Where applicable, Individuals must comply with the Government Ethics Laws and Standards of Conduct with regard to conflicting financial interests and other appearances of bias. For non-employees, procedures should be established to identify and appropriately deal with individual conflicts of financial interest or other potential sources of bias.
- Individuals with a clear vested (or conflict of) interest in a topic can better serve as resources to a panel to assure that all evidence has been presented and considered rather than as voting panelists.
- Panelists should consist of individuals with methodologic expertise in assessing scientific evidence (epidemiologists, statisticians, and economists), and representatives of likely primary users of the guidelines (medical specialists, primary care and public health practitioners, and administrators).
- Consumer representation should be considered.
- The procedures and criteria for selecting panel members should be documented.

Multidisciplinary panels offer the advantage of providing more balanced perspectives. They should also represent the range of culturally diverse populations which will likely be users of the guidelines. While bias is a potential problem with homogeneous panels, for instance of physicians, heterogeneous panels with ethicists, consumers, and others raise concerns about their technical ability to interpret and use epidemiologic and other scientific evidence.

Individuals with a stated opinion (published articles, editorials, or speeches) on a topic may also have difficulty objectively reviewing evidence which may conflict with their stated perspective. While it is usually difficult to convene a multidisciplinary panel to deal with a wide

Approaches to Selecting Panels

variety of topics without having individuals with stated positions on issues before the panel, individual panel members can excuse themselves from voting on issues where they have previously established positions.

Federally chartered advisory committees (with or without rotating membership), e.g. ACIP, provide continuity of methodology and the knowledge base and are most suitable when an ongoing need is anticipated. For recommendations requiring very specialized knowledge, such as laboratory procedures for diagnostic testing, a highly technical group, possibly consisting of internal CDC staff alone or in consultation with outside experts, may be desirable. When timeliness is imperative, such as for an international outbreak, it may suffice to update previous recommendations or to adapt them based on recent information. Under similar circumstances it may be necessary to disseminate recommendations without extensive external review, or to issue interim guidance which later can be made final after more intensive review and revision. For more complex subjects or where synthesis of the literature is required, a panel is more suitable.

Consideration must also be given to the intended users of the guidelines. Inclusion of representatives from interest groups of likely users, either as voting or non-voting liaison members, should enhance the likelihood that the users will accept and implement the recommendations because they will have confidence that their perspectives were considered during the panel's deliberations. Nonetheless, it is often impractical to include all interested parties on most panels.

To facilitate group processes, panels are often limited to 10 to 20 members. Nonetheless, the complexity of the tasks (topic selection, development of rules of evidence, development of evidence-based models, identification of alternative technologies, review of literature, analysis and synthesis of the evidence, evaluation of effectiveness, costs, and safety, identification of research needs, and comprehension of policy-related issues) may require skills which cannot be fully captured in a small group of individuals. Moreover, should panels be convened to write one or a limited number of guidelines, the expertise gained as part of the guidelines-development processes will be lost.

The OTA has suggested the creation of expert teams to support guideline panels. These teams could study alternative strategies for performing the panel's tasks and develop a recommended set of methods based on empirical experience. CDC staff often provide staff support, though the creation of clearly defined teams is not common. Another model for support would be formation of a center, e.g., in an academic institution, which can perform staff support (convene meetings, perform literature review, conduct statistical analyses, and provide editorial and clerical support), or a contract to provide selected support functions.

The Cochrane Collaboration, an international network of individuals and institutions committed to preparing, maintaining, and disseminat-

ing systematic reviews of the effects of health care, employs a Collaborative Review Group (CRG) to prepare a systematic review of randomized controlled trials (RCT) and other evidence relating to a particular health care maneuver. The CRG starts with a group of self-nominated individuals (health professionals, methodologists, and consumers) who share a common interest in an issue in the effectiveness of health care. As an outgrowth of one or more exploratory meetings, one or several teams decide to take on, for the rest of their careers, the task of preparing and maintaining systemic reviews of RCTs on that issue (Sackett, 1994).

As described in earlier sections of the document, guidelines are written or at least approved by a recognized group or committee which has supervision over the guidelines development process. Some are federally chartered advisory committees (ACIP, HICPAC); others are less formal and may be composed of CDC staff or selected experts. Some are collaborative activities between outside organizations (American Lung Association and American Thoracic Society) and CDC operating units e.g., the Division of Tuberculosis Elimination. Panels may be standing committees, may rotate membership, or be convened ad hoc.

Because of the variety of approaches to selecting expert panels to develop CDC-sponsored guidelines, greater effort is needed to assure diversity of expertise, interests, and cultural sensitivities among participants.

Current Practice at CDC

This page intentionally left blank.

Defining the Scope of a Guideline

In defining the scope, guideline developers explicitly delimit, and usually narrow, a broad topic. In doing so they must focus on the preventive, diagnostic, treatment, or rehabilitative issues which are (1) of greatest interest to practitioners, the public, and other users, (2) most amenable to clarification by means of systematic assessment and synthesis of scientific evidence, (3) in greatest need of having guidelines because of conflicting claims of effectiveness in published reviews, and (4) capable of being fully explored and resolved into clear and specific advice within the limitations of time and resources (Field, 1995).

Explicit answers to the following questions can help establish the scope or topic boundary of the guideline.

Points to Consider

- What target population (age, gender, ethnicity or race, socioeconomic status, etc.) and service settings (state or local health agency, community, home, workplace, hospital, ambulatory clinic) will the guideline encompass?
- What population-specific outcome measures (risk factor prevalence, morbidity, mortality, quality of life) or individual health condition (healthy, at-risk and asymptomatic, symptomatic, diseased) will the guideline aim to improve?
- What intervention (population- or individual-based), essential public health service (see page 108 of glossary) or provider practice will the guideline aim to improve?
- What characteristics (safety, effectiveness, appropriateness, cost, feasibility of implementation, or patient preferences) of the intervention or practice options should be used to define good-, better, and best-practices?
- Who are the guideline users (public, legislator, agency manager, lay advocate, methodologist, clinician) whose knowledge, practices, behavior, or decisions will the guideline attempt to address?
- How much time, staff, and other resources are available to complete the task?
- What is the scope of existing guidelines on the same topic and what information gap would be filled by the proposed guideline?
- What new work has been done since publication of the existing guidelines?

Recommendations

These recommendations emphasize the role of CDC staff during the critical phase of planning, financing, and preparing for the guideline development effort prior to the convening of an expert panel or other form of deliberative group. Because empirical evidence suggesting that any of the approaches to defining the scope of a guideline described below is more or less effective than another is not available, these recommendations reflect a narrative synthesis of the professional opinions of the experts whose work is cited.

- CDC staff should prepare explicit answers to the questions in the Points to Consider as early in the guideline development planning process as is feasible.
- Using preliminary answers to those questions, guideline planning staff should conduct a preliminary assessment of the amount and quality of the scientific evidence of effective practices and implementation approaches that is likely to be available under alternative scenarios of a guideline whose scope varies from narrower to broader.
- Preliminary decisions taken by the guideline planning staff before an official panel is convened should be revisited as soon as possible after the panel leadership is appointed and the members are confirmed.
- When the use of a high-volume practice of unproven effectiveness is widespread, relevant guidelines should address issues of the safety and effectiveness of the practice. For example, guidelines on the prevention and control of prostate cancer should address the use of the prostate specific antigen (PSA) test.
- Issues related to the implementation of guidelines (cost, cost-effectiveness, staffing, insurance coverage, and patient preferences) should be taken into account whether or not these issues are used to define best-practices.

Approaches to Defining the Scope of a Guideline

Public and private agencies which sponsor the development of guidelines have used several alternative approaches to defining the scope of a guideline. These approaches can be put into three broad categories based on the primary source of influence on decisions about the scope of a guideline. The categories are (1) legislative approach, (2) expert panel approach, and (3) internal staff approach. More often than not, however, elements of each approach are combined in the process of defining the scope of a guideline. These three approaches are illustrated by reference to the work of selected governmental agencies and private groups which sponsor guidelines.

Legislative Approach

The general scope of guidelines sponsored by the Agency for Health Care Policy and Research (AHCPR) is prescribed in the agency's

authorizing legislation. For example, the law requires special attention to meeting the needs of the Medicare program. However, the specific scope of each guideline sponsored by the Agency has been determined jointly by an appointed expert panel and Agency staff using a process already described. Schriger, one of two participating methodologists, described the procedure used by the panel on Low Back Pain to define the scope of its guidelines. The time between the selection of the panel and the first meeting was used by the methodologists, panel chairperson, and agency staff to negotiate a panel budget and appropriate scope of work which could be accomplished by means of four panel meetings, each lasting 1.5 to 3 days, during a 12-month period. Final decisions about the elements of guideline scope were made during the course of the first meeting. The methodologist's most important educational role with respect to guideline scope was to restrain tendencies by both the Agency staff and the panel chair to expand the scope of the guideline. (Schriger, 1994, p 118).

Most guideline development efforts depend on the expert panel, which is convened to judge the quality of the science and to compose the recommendations, also to define the scope of the guideline at the beginning of deliberations. This process is exemplified by the guideline development efforts of the National Heart, Lung, and Blood Institute (NHLBI), the USPSTF, the Canadian Community Health Practice Guidelines (CHPG) Working Group, and the Council on Linkages Between Academia and Public Health Practice, Guideline Development Project for Public Health Practice (GDPPHP). Each of these groups has used a group of key questions to narrow the scope of the guideline topic.

This approach is exemplified by the Consensus Development Conference Program of the NIH, Office of Medical Applications of Research (OMAR). It uses a planning committee made up of two to three non-government researchers, an agency staff person, and a staff person from the sponsoring institute(s). The planning group identifies four-to-six key questions to be answered at the conference. The questions usually relate to efficacy, risks, clinical applications, and avenues for future research. The wording of the questions frames the scope of the resulting consensus statements.

Historical information about defining the scope of CDC-sponsored guidelines is not well documented. However, in April 1995, CIO staff identified 96 guidelines which met a Public Health Service (PHS) case definition, i.e., they were produced and/or funded by PHS, published in or after 1989, and intended to influence care to individuals (including prevention, screening, diagnosis, and treatment—irrespective of the setting in which the service is delivered). Of the 96 PHS-defined “guidelines intended to influence care to individuals,” 53% were primarily intended for use by practitioners, 30% for use by programs, 11% for use by systems which cut across programs and institutions,

Expert Panel Approach

Internal Staff Approach

Current Practice at CDC

and 5% for use by patients and/or parents. Clinical conditions were the focus of 90% and procedures were the focus of 10%. Among dimensions of intervention that were not mutually exclusive, 77% emphasized primary prevention, 43% screening, 44% diagnosis, 40% treatment, and 2% rehabilitation. Presumably, current guideline development efforts at CDC reflect this variability of approaches to defining the scope of guidelines.

Clarifying the Method and Analytic Framework

After the CDC program (usually) or expert panel has chosen a guideline topic and defined its scope, the participants should choose a guideline development method, define the analytic framework, collect and assess evidence of effectiveness (or lack thereof) of candidate practices, and compose recommendations that will help potential users make decisions about adopting (or discontinuing) those practices. A guideline development method is a set of rules of procedure for collecting evidence of effectiveness of a health practice and for making group decisions about the quality and sufficiency of that evidence in supporting specific practice recommendations.

Woolf has delineated four categories of guideline development methods. The categories are (1) **informal consensus**, (2) **formal consensus**, (3) **evidence-based**, and (4) **explicit**—an evidence-based method which incorporates mathematical modeling. Experts are more likely to consider evidence-based (and explicit) methods valid in content, reproducible, and scientifically rigorous than consensus methods. In addition, evidence-based (and explicit) methods are more likely than consensus methods to be costly, difficult to execute, and require more data and time to complete. Finally, critics of the more rigorous of guideline development methods point to the frequency with which the absence of acceptable evidence leads to neutral recommendations neither for nor against the proposed preventive practice. Most guideline development efforts use a combination of methods. Regardless of the category of development method chosen, the expert panel must explicitly define the guideline's analytic framework (Woolf, 1992).

Defining the **analytic framework** refers to describing or illustrating the chain of causal reasoning which links the recommended health practice to the desirable health outcomes in a defined individual or population by means of credible evidence of effectiveness. The analytic framework is an effective conceptual tool for planning, conducting, and communicating the results of the guideline development process to potential users. It helps developers and potential users understand which parts of the supporting rationale for the practice recommendations are based on empirical evidence and which components are based on theory, expert opinion, conventional standards of practice, or other potential sources of influence on the clinical and population-based practices (Woolf, 1994).

The analytic framework also makes it easier to compose the supporting rationale for each practice recommendation. A persuasive supporting

rationale for a practice recommendation would summarize (1) the benefits, harms, and other outcomes that were considered; (2) why the outcomes were considered important; (3) assumptions about the relationships between categories of benefits, harms, and outcomes; (4) the types of evidence ultimately used to support each recommendation; (5) and as one type of evidence, the estimated effects of the practices recommended.

Points to Consider

Explicit answers to the following questions about the proposed guidelines can help construct the analytic framework:

- What professional practices or intervention strategies are being considered for inclusion in the guideline?
- What categories of health outcomes (ultimate, intermediate, and surrogate) are the guidelines intended to influence?
- What specific guideline development methods (explicit, evidence-based, formal consensus, informal consensus) is most appropriate for the proposed topic?
 - What types of evidence of effectiveness are available to support the professional practices being considered?
 - How will the quality, quantity, and relevance of the supporting evidence of effectiveness be assessed?
 - How will the hierarchy (or strength) of the evidence relate to the strength of a recommendation for or against a proposed intervention?
- What assumptions about proposed practices, health outcomes, and their causal relationships are offered without scientific proof?
- What form of exposition (narrative or graphic—flow chart, influence diagram, decision tree, or clinical algorithm) is most appropriate for describing the logical link between the practice recommendations, evidence of effectiveness, and desirable health outcomes?

Recommendations

The following recommendations for defining the analytic framework for CDC-sponsored guidelines apply to all categories of guideline development methods. These recommendations emphasize the consensus among experts that guideline development methods should be evidence-based or explicit, whenever possible.

- Guidelines about preventive interventions should be based on a consideration of relative benefits, harms, and costs, whenever possible.

- Many experts assert that evidence-based methods lead to guidelines that are more valid (face or content validity) and credible than those resulting from consensus methods (informal or formal) (Field and Lohr, 1992).
- Evidence-based methods of guideline development are complex, time-consuming, and depend on skills that potential participants may not possess at the start of the process. Thus, when undertaking evidence-based methods, all participants must be suitably trained in the skills needed to carry out their particular tasks as early in the guideline development planning process as is feasible. The services of a guideline development methodologist should be available to the panel.
- Just as every research study or epidemiologic investigation should be guided by a written analytic plan, so too every guideline development effort should be guided by a written analytic framework (narrative, algorithm, spreadsheet, etc.).
- The analytic framework should be defined by providing explicit answers to the questions listed in the preceding Points to Consider.
- CDC staff should consider the rules and regulations of other agencies such as the Food and Drug Administration (FDA). Vaccines, biologicals, drugs, and medical devices must meet statutory requirements regarding their safety and efficacy before their use is approved. FDA approval is based on a biostatistical regulatory model, a special form of evidence-based model. Because guidelines on the use of these regulated products are based on an analytic framework which is, in part, prescribed by law, CDC staff and expert panels working in this area should receive special training in the use of this model (O'Neill, 1994).

The analytic framework often is illustrated graphically (Figure 1) and serves as a useful guide for reviewing evidence of effectiveness, developing recommendations, and explaining the supporting rationale. For example, an expert panel charged with developing guidelines for the use of estrogen replacement therapy to reduce morbidity and mortality (ultimate outcomes) among post-menopausal women might choose to use an evidence-based methodology. The intermediate outcomes of interest might include reduction (of incidence or prevalence) of heart disease, osteoporotic fractures, and menopausal symptoms. The types of admissible evidence of effectiveness might come from randomized con-

Approaches to Method and Analytic Framework

Examples From Other Agencies

trolled trials; prospective-, retrospective-, and cross-sectional observational studies; expert opinion; and current practice.

Criteria for assessing the strength of the evidence supporting a particular recommendation might include indicators of quality (e.g., study design), quantity (e.g., effect size), and relevance (e.g., external validity or ability to generalize). Criteria for assessing the appropriateness of a recommended practice in a particular setting might also include indicators of cost-effectiveness, ease of implementation, and preferences of affected providers, clients, and communities (Woolf, 1994).

Public and private agencies which sponsor the development of guidelines have used several alternative approaches to defining the analytic framework of a guideline.

Examples of alternative approaches to defining the analytic framework of a guideline are described below. Although the examples are grouped by Woolf's categories of methods, for convenience, the emphasis of the description is on the analytic frameworks rather than the overall method.

Informal Consensus

The analytic framework of informal consensus methods usually cannot be discerned from the guidelines. Using this method, experts decide on what practices to recommend after informal discussions. The recommendations are often not accompanied by explanations of the causal reasoning which links the recommended professional behavior, credible evidence of effectiveness, and the desirable health outcomes. Similarly, it may not be clear which parts of the supporting rationale for a recommendation are based on empirical evidence and which components are based on theory, expert opinion, conventional standards of practice, or other potential sources of influence on the clinical and population-based practices of the intended guideline users. Because of its simplicity and speed, this is the most common approach to guideline development. Unfortunately, it is also the least valid (Woolf, 1992).

Formal Consensus

The analytic framework of formal consensus methods is usually defined by a series of key questions. This analytic approach to developing guidelines was pioneered by the Consensus Development Conference Program of the National Institute for Health (NIH). The Consensus Development Conference Program of the National Institute for Health (NIH) aims to identify and disseminate clinically relevant findings emerging from NIH research, often during initial technology diffusion and before guidelines are published. Consensus statements are either condition specific (e.g. melanoma or panic disorder) or technology specific (e.g., dialysis or physical activity). The analytic approach to a consensus conference is framed by a series of four to six questions

related to efficacy, risks, clinical applications, and future research. For example, the key questions for a December 1995 conference on physical activity and cardiovascular health were as follows:(NIH, 1995)

- What is the health burden of a sedentary lifestyle on the population?
- What type, what intensity, and what quantity of physical activity is important to prevent cardiovascular disease?
- What are the benefits and risks of different types of physical activity for people with cardiovascular disease?
- What are the successful approaches to adopting and maintaining a physically active lifestyle?
- What are the important questions for future research?

Answers to the framing questions are arrived at by an expert panel of 20 to 30 members who make public presentations on the science and an executive panel of 9 to 16 members who draft the consensus statements. Consensus statements do not usually include references to the literature, nor the underlying rationale or evidence behind any recommendations or conclusions (US Congress, OTA, 1994).

The NIH consensus process has been vigorously criticized because of the explicit lack of rationales and evidence which support the statements. Improvements in the process have been recommended (Field and Lohr, 1992).

Evidence-Based, Including Explicit Approaches

Evidence-based approaches to guideline development depend on analytic frameworks that aim to explicitly link practice recommendations to the underlying scientific evidence of effectiveness. The explicit approach, a subcategory of the evidence-based approach, is distinguished by its reliance on mathematical modeling and other formal analytic methods to generate estimates of the probability of occurrence of specific benefits, harms, and costs of alternative practices. This approach also uses a “balance sheet” to display the benefits, harms, and costs of each alternative practice along with the content and source of necessary assumptions which are not scientifically disprovable (Woolf, 1992).

- **The Agency for Health Care Policy and Research (AHCPR)** sponsors guidelines which focus on the diagnosis and management of clinical conditions e.g., congestive heart failure and low back pain. AHCPR guidelines are developed by independent expert panels using evidence-based methods. Panels use “evidence tables” to summarize important aspects such as the design of relevant studies; and most panels describe the strength of the supporting evidence for conclusions and recommendations. Some AHCPR panels have

used variations of Woolf's "evidence model" while others have used Eddy's explicit approach (Eddy, 1990). Both of these guideline development methods require that the analytic framework be defined explicitly.

AHCPR-sponsored guideline recommendations are often illustrated using a clinical algorithm or flow chart that shows recommended steps in diagnosis or management. Algorithms help guideline developers to specify the appropriate indications for a particular diagnostic or management strategy and to convey the scope of the guideline to the user at a glance. Two major criticisms of clinical algorithms include inflexibility and questionable validity. To increase clinical flexibility, AHCPR-sponsored clinical algorithms include special counseling and decision nodes showing where major preference-dependent decisions occur. To increase the clinical validity of their algorithms, some AHCPR panels have used annotations which link recommendations and their expected outcomes to the literature through evidence tables that summarize the relevant studies (Woolf, 1992; US Congress, OTA, 1994).

- **The U.S. Preventive Services Task Force (USPSTF)** develops evidence-based practice guidelines for preventive care which can be appropriately delivered by primary care practitioners during a periodic health examination. Like the AHCPR guidelines, the USPSTF guidelines have used variations of Woolf's "evidence model" and Eddy's explicit approach. Both of these guideline development methods require that the analytic framework be defined explicitly (US Congress, OTA, 1994). Specifically, the USPSTF has refined a method for reviewing evidence of effectiveness of a proposed intervention. The available evidence for or against the effectiveness of the proposed intervention is classified into one of five levels of decreasing quality (I, II-1, II-2, II-3, III); and the resulting recommendation for or against routine use of the intervention is classified into one of four levels of strength (A,B,C,D)(Table 2)(USPSTF, 1996). CDC staff who support the work of HICPAC have expressed some discomfort with the USPSTF. They would prefer to use categories A, B, or C to describe the strength of all recommendations including advice against the use of an ineffective practice (Bill Martone, Personal Communication).
- **The Canadian Community Health Practice Guidelines (CHPG) Working Group** developed practice recommendations for restaurant inspections, STD partner notification, and immunization delivery, which served as prototypes to develop a standardized approach. The analytic framework of the standardized approach was defined by an intervention-outcome grid in the form of a spreadsheet. Each row of the grid represents a specific intervention, each column represents an outcome (benefit, harm, or cost) and each cell represents the attributable effects of the intervention

on the outcome. The grid specified at the beginning of the process is used to specify selection criteria for the evidence to be reviewed. It is modified as a result of the evidence discovered during the process, and it is used at the end of the process to summarize the relationship of the interventions, outcomes, and evidence of effectiveness.

Each resulting recommendation was reported along with its level of evidence from comparative studies, and justification or basis. Recommendations were classified as “plus” if evidence from high quality comparative studies suggested that the intervention should be considered for inclusion in routine public health practice; “minus” if the evidence suggested that it should not; and “questionable” if the evidence from comparative studies was lacking or contradictory and other grounds for a recommendation are unpersuasive or conflicting (Corber, 1994).

- **The Council on Linkages Between Academia and Public Health Practice** issued a report on the feasibility and benefits of practice guidelines for public health in October 1995. The Guideline Development Project for Public Health Practice (GDPPHP) tested a guideline development model in each of four health problems: (1) pre-school immunization, (2) lead poisoning prevention, (3) tuberculosis treatment completion, and (4) ischemic heart disease prevention.

The project leaders and panel chairs developed a series of Critical Questions to guide the search for scientific evidence related to the 10 essential public health services (see page 108) in each program area. Reflecting on their experiences with the “essential services/critical questions model” for defining the analytic frameworks for each program area, the panelists concluded that the model may be useful for gathering evidence and also for conceptualizing and focusing future guidelines for public health practice. The panelists who participated in this feasibility study recommended that the “critical questions” of the future should be (1) based on issues of practice as well as theory, and (2) developed with the involvement of major stakeholders at the beginning of the process (Council on Linkages between Academia and Public Health, 1995).

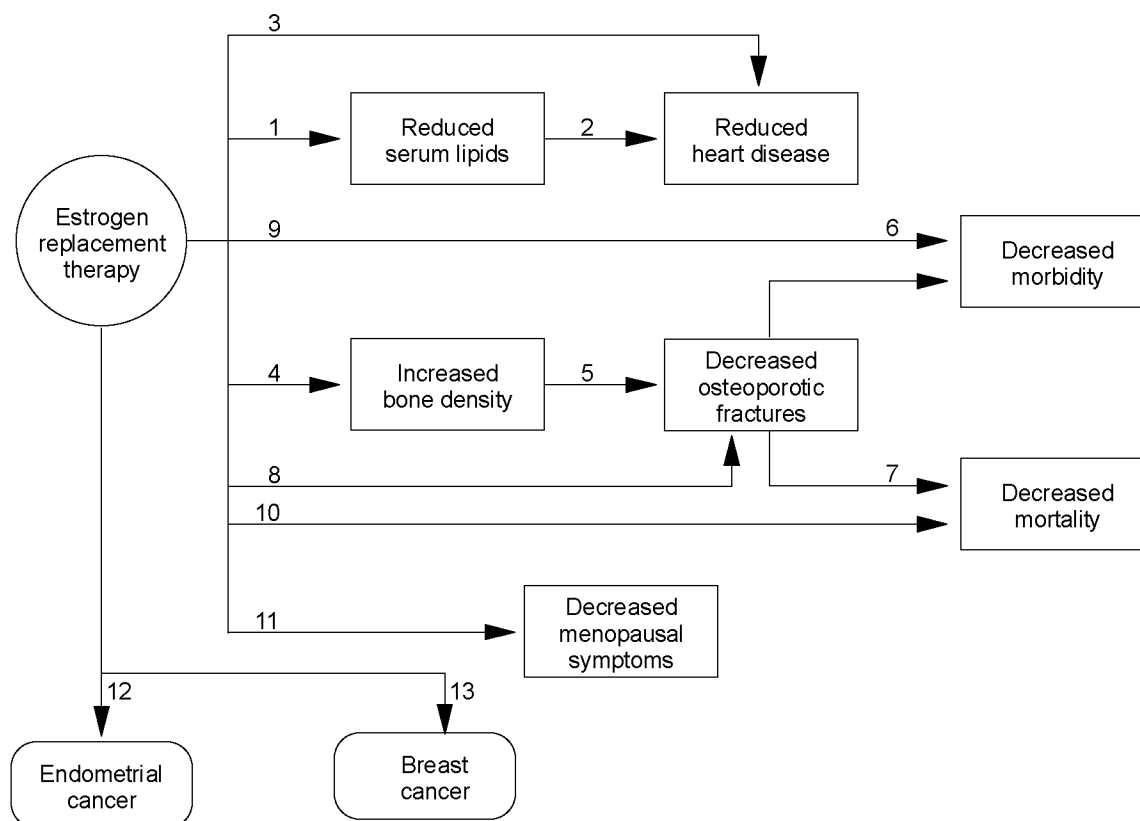
- **The FDA Biostatistical Regulatory Model** aids the implementation of regulations concerning the standards for evidence of efficacy and safety needed for the approval of new drugs, vaccines, medical devices, and biologicals. For example, FDA scientists at the Center for Drug Evaluation and Research use the model to evaluate the quality, documentation, and adequacy of clinical research evidence submitted by pharmaceutical sponsors in support of applications for approval of new drugs. Elements of the model including standards of evidence, documentation of evidence, and the review process are clearly spelled out in sections of the Code of Federal Regulations(CFR). The standards provide the framework

within which the agency approves or disapproves a new drug application which claims a specific outcome when the drug is used as prescribed in a specific target population (O'Neill, 1994).

Current Practice at CDC

Current guideline development efforts at the CDC employ one or more of the categories of guideline development methods described above. In response to a 1995 PHS survey, CDC staff identified 96 CDC guidelines which were “intended to influence care to individuals” and were published in or after 1989. The respondents reported that 26% of the guidelines used a consensus method (informal or formal), 29% used an evidence-based method, and 40% used some approach in which neither a consensus nor evidence-based method was predominant (PHS Sponsored Guidelines Report Outline, 1995).

Figure 1. Analytic framework for estrogen replacement therapy*



Note: Numbers represent linkages.

- Link 1:** Evidence that estrogen reduces serum lipid levels?
- Link 2:** Evidence that lowering serum lipid levels reduces the incidence of heart disease?
- Link 3:** Direct evidence that estrogen reduces the risk of heart disease?
- Link 4:** Evidence that estrogen increases bone mineral content?
- Link 5:** Evidence that increased bone mineral content is associated with decreased risk of fractures?
- Link 6:** Evidence that fractures are associated with increased morbidity?
- Link 7:** Evidence that fractures are associated with increased mortality?
- Link 8:** Direct evidence that estrogen reduces the incidence of fractures?
- Link 9:** Direct evidence that estrogen reduces morbidity?
- Link 10:** Direct evidence that estrogen reduces mortality?
- Link 11:** Evidence that estrogen reduces menopausal symptoms?
- Link 12:** Evidence that estrogen causes endometrial cancer?
- Link 13:** Evidence that estrogen causes breast cancer?

* Reproduced from (Woolf, 1994)

Table 2. Rating Scales: From USPSTF

Quality of Evidence (of Effectiveness of the Intervention)*

- I:** Evidence obtained from at least one properly randomized controlled trial.
- II-1:** Evidence obtained from well-designed controlled trials without randomization.
- II-2:** Evidence obtained from well-designed cohort or case-control analytic studies, preferably from more than one center or research group.
- II-3:** Evidence obtained from multiple time series with or without the intervention. Dramatic results in uncontrolled experiments (such as results of the introduction of penicillin treatment in the 1940s) could also be regarded as this type of evidence.
- III:** Opinions of respected authorities, based on clinical experience; descriptive studies and case reports; or reports of expert committees.

Strength of recommendations (for or against the intervention)**

- A:** There is **good evidence to support** the recommendation that the condition be specifically considered in a periodic health examination.
- B:** There is **fair evidence to support** the recommendation that the condition be specifically considered in a periodic health examination.
- C:** There is **insufficient evidence** to recommend **for or against** the inclusion of the condition in a periodic health examination, but recommendations may be made on other grounds.
- D:** There is **fair evidence to support** the recommendation that the condition be excluded from consideration in a periodic health examination.
- E:** There is **good evidence to support** the recommendation that the condition be excluded from consideration in a periodic health examination.

* Well-designed and conducted meta-analyses were also considered and were graded according to the quality of the studies on which they were based.

** Exact correlation between level of evidence and strength of recommendation did not occur. Level I evidence did not necessarily lead to a grade A recommendation, nor did Level III evidence necessarily lead to a grade D recommendation.

Identifying and Synthesizing the Evidence

Recommendations regarding public health practice should be based on the combined weight of the evidence from available scientific reports. To accomplish this objective, the available literature must be identified, retrieved and synthesized to draw scientifically valid and practically useful conclusions. The National Library of Medicine (NLM) has suggested a systematic process for identifying, retrieving, and managing literature that forms the basis of a practice guideline (Auston, 1994). The six steps are:

- Advance planning to define goals, tasks, staffing, and computer software needs.
- Preliminary searching to estimate content and volume of evidence expected.
- Comprehensive searching based on explicit statements of topic scope, selection criteria, sources, and search strategies.
- Literature management of the retrieved citations in the form of a bibliographic database.
- Document retrieval and archiving of copies of relevant documents.
- Final bibliography preparation to support panel deliberations and citations for evidence tables, guideline text, and appendixes.

The remainder of this section of the guide focuses on issues related to synthesizing evidence of effectiveness of a candidate practice guideline. Research synthesis can be done qualitatively based on individual opinion or group consensus or more rigorously using the tools of systematic reviews or meta-analysis. The least rigorous approach to synthesizing evidence is through expert opinion, often obtained through committees. In its most basic form, this approach is qualitative without formal data collection or analysis. More rigorous, narrative literature reviews are conducted by experts in the field and written as expert opinion papers in journals and text books, for example. For these, however, the lack of systematic rules for acquiring or consolidating the evidence has led to questions about the validity of such reviews (Antman, 1992).

Systematic literature reviews have been introduced as an appropriate method of data collection and synthesizing evidence (Thacker, 1988; Pettiti, 1994). These reviews are based on rules to identify studies and collate information from studies. Such reviews are often the initial step

in meta-analyses where the data are synthesized quantitatively to give a numerical estimate of effect.

Despite the strengths and weaknesses of the alternative methodologies for combining evidence, systematic reviews and meta-analyses remain the best available tools for synthesizing evidence, particularly when results from independent studies are inconsistent or primary data are not available for pooling.

Consensus on interpreting data requires some form of group judgment before public health guidelines are specified. Group judgment efforts for evaluating practices and procedures must bridge gaps and resolve disparity among research findings, define the state of the art, and establish public health policies. Expert panels must integrate, interpret, and weigh evidence, experiences, beliefs, and values, and then formulate guidelines and recommendations. Evidence may consist of systematic summaries, such as one finds in a meta-analysis, or more commonly a sparse patchwork of contradictory research results of varying quality.

In the final synthesis of evidence, the results of the systematic reviews and meta-analyses are weighed together with other evidence (e.g., cost and ethical considerations) by experts in order to formulate policy and make recommendations.

The best group judgment efforts delineate their assumptions when confronted with inconsistencies, contradictions, and gaps in research. These group efforts provide a forum for participants to learn from each other and serve as effective means to disseminate findings through key opinion leaders within the group.

Points to Consider

Systematic reviews of the literature are an important basis for developing CDC guidelines and always should be used. Meta-analysis or quantitative synthesis, a form of systematic review, should be used more regularly when appropriate. It encourages systematic thought concerning methods, outcomes, categorizations, populations, and interventions, and provides a way to synthesize evidence. Answers to the following questions can help determine the appropriateness of meta-analysis:

- Are results from multiple independent studies inconsistent?
- Do multiple studies suggest an effect but no single study has sufficient power to demonstrate statistical significance?
- Is a quantitative synthesis needed for policymaking?
- Is it necessary to study multiple population characteristics (occupation, age, gender, race, ethnicity, medical condition, or geography)?

- Data synthesis including a systematic review of the literature, and where relevant and feasible meta-analysis, should be the basis for development of all CDC guidelines.
- The questions to be addressed by the panel should be specific and manageable. Panelists should be involved in specifying the questions to be addressed and responsibilities for task and project format.
- Panelists should be provided with the most comprehensive scientific data possible. A summary description of the available studies should be provided and cited in the final panel statement.
- **Process methodology, facts, assumptions, estimates, criteria for findings, and rationale** should be documented. Findings should include estimates of **outcomes expected** if the panel's recommendations are followed.

Recommendations

Approaches to Synthesizing Evidence

Traditional

In a traditional narrative review of the medical literature, a subject-matter expert reviews studies, decides which are relevant to a particular topic, and highlights findings in terms of results and, to a lesser degree, methodology. The potential limitations of this, or any approach to a literature review, include (a) bias due to reporting and publication policies, (b) the absence in published studies of specific data needed for the review, (c) investigator bias due to subjective criteria for including studies, (d) the uneven quality of the primary data, (e) improper attention to statistical methods, and (f) biased interpretation of outcome. (Thacker, 1988).

Meta-analysis is not immune to the potential pitfalls noted for the traditional narrative review. In addition, the approach may not be feasible for developing guidelines that address complex topics such as nosocomial infection control. Nevertheless, the technique makes decisions apparent and makes explicit the potential impact of problems (e.g., bias) on the results and interpretation. Meta-analysis encourages systematic thought concerning methods, outcomes, categorizations, populations, and interventions and provides a way to synthesize evidence. Further, the method offers a mechanism for estimating the magnitude of effect in terms of a statistical measure (e.g., an odds ratio or relative risk) and an assessment of its significance.

Meta-Analysis

The combination of data from several studies increases both the statistical power and ability to generalize, thereby enabling the researcher to assess more completely the impact of a procedure or variable. Quantitative measures across subgroups of studies can provide insight into the

nature of relationships between variables, offering mechanisms for detecting and exploring apparent contradictions and results. Finally, a systematic review should be less subjective than a narrative review, and although the procedure may not decrease investigator bias, it enables the reader to understand clearly how and why conclusions were drawn.

Systematic rules for conducting a meta-analysis include an explicit description of methodology so that results can be interpreted in light of any biases or limitations. (See Points to Consider for ways to help determine if meta-analysis is appropriate.)

Several approaches to meta-analysis share the same basic steps listed below:

- A clear statement of the problem and an explicit statement of the hypothesis to be tested.
- A clearly defined statement of inclusion and exclusion criteria for admission of studies.
- A methodology for locating research studies.
- The classification and coding of study units to be combined in the meta-analysis.
- A quantitative measurement of study characteristics on a common scale.
- A quality assessment of the methods used in the studies.
- Analysis and interpretation that include determining the homogeneity of the data and, where appropriate, combining study results.
- The use of appropriate statistical models.
- An interpretation of results.
- The reporting of the results (Thacker, 1988; Pettiti, 1994).

A particular problem facing those who use meta-analysis for policy formulation is the fact that studies in human populations (whether a single patient or a community) form a continuum of knowledge as new studies are applied to similar populations. This fluid nature of ever-increasing data can be addressed by a cumulative meta-analysis in which results are updated as new information becomes available (Lau, 1992). This technique is particularly useful for studies (particularly trials) where the protocol may be modified in the process of study completion or where preliminary results may be disseminated (Henderson, 1995).

Statistical issues related to combining data from multiple sources in meta-analysis are the subjects of ongoing research (Antman, 1992; Steinberg, 1991). The most important statistical issues relate to the question of which studies should be combined. Meta-analyses appropriate

for the public will usually be a combination of randomized, controlled clinical trials (RCTs) and/or epidemiologic studies. In general, RCTs provide more compelling evidence than do cohort studies, and, in turn, cohort studies often provide better evidence than do case-control studies. In some situations, however, cohort studies may not be feasible (rare disease or disease with long latency) or may be less helpful than case-control studies (technology assessment). Cross-sectional studies and case series provide less evidence for etiologic reasoning (Hedges, 1985). Thus, when feasible, meta-analyses in public health should use RCTs. However, because of the small number of RCTs available to test hypotheses of interest in public health or the time lag until results are available, the combining of data from cohort and case-control studies is sometimes desirable (Greenland, 1994). In such circumstances, the statistical and epidemiologic issues are more complex and challenging to the meta-analyst (Hedges, 1985; Greenland, 1987; Greenland, 1994).

Like traditional narrative literature reviews, the usefulness of meta-analysis depends, to a large extent, on the quality of the studies that are synthesized. Methodologic concerns about meta-analysis include limitations in the quality of the primary data and both selection and investigator bias in interpretation, but this is true for any literature review. Concerns more specific to meta-analysis include the appropriateness of combining data across studies, appropriate use of statistical methods, variability between studies, development of appropriate models to measure such variability, and the role of assessing study quality (Greenland, 1987; Greenland, 1994; Eysenck, 1994). Creative approaches to addressing these concerns are being developed (Sterling, 1995; Emerson, 1990; Oxman, 1991). Finally, in controlled trials in which whole communities are randomized, a large enough number of eligible studies may never be available to support a meta-analysis.

A serious concern about quantitative literature reviews is that the reported objectivity may be more appearance than substance; no approach to synthesizing information can eliminate investigator bias entirely. Both critics and advocates of meta-analysis recognize that an unwarranted sense of scientific validity, rather than a more accurate understanding, may result from quantification.

This page intentionally left blank.

Aiding Group Interaction and Decision Making

Reaching a conclusion can be a difficult task for a group engaged in decision making, and in a multi-faceted decision like the development of a scientific guideline, it can be even more difficult. Participants' diversity in terms of expectations, expertise, and stake in outcome, as well as their cultural diversity (both ethnic cultures and organizational cultures), contribute to the complexity of group process. The techniques described in this section can aid groups in producing high-quality, science-based guidelines by allowing groups to take advantage of the contributions of all participants. The first step in the process of planning for group activities is to determine whether assembling a group is the most effective way to accomplish a particular task. Groups should not be encumbered with tasks that individuals can accomplish more efficiently. Although groups are better at generating ideas, individuals are better at research, analysis, and careful crafting of language (Moore, 1994). Also keep in mind that the type of decision making process used may depend on the group structure, and whether it is legally appropriate for the group to reach consensus.

Group interaction methods (or group process methods) improve productivity of meetings by helping groups focus on tasks and produce outcomes. Expert facilitators say that group interaction techniques improve the quality of virtually any kind of meeting, but especially those between people without a history of working together. For maximum usefulness, the group interaction method must be appropriately matched to the purpose of the meeting and its participants (See Table 3). Although there is no evidence (yet) to suggest the superiority of any particular method in guideline development, group interaction methods generally enhance group process and decision making (US Congress, OTA, 1994).

In developing guidelines, group interaction methods can be used to (1) organize and manage the process of guidelines development, (2) structure group meetings to balance power and manage bias by reducing the potential for the most powerful or vocal group members to dominate the process, (3) provide a framework for group decision making, and (4) help groups generate more ideas than if they worked individually while moving the group toward decision making. A thoughtfully prepared plan for group interaction supports the intent to develop guidelines in a logical, systematic and accountable manner. Using techniques shown to be effective adds to the credibility of the decisions made and facilitates effective and efficient working groups. Group interaction methodologies also provide documentation of the

development process which adds clarity and explicitness regarding group decision making.

Points to Consider

- Is it legally appropriate for this type of group to reach consensus, or is the goal to obtain individual input?
- Are CDC staff and the panelists trained in the appropriate use of group interaction techniques?
- Are groups being assigned tasks that can be accomplished more efficiently and effectively by group process rather than by individuals?
- Considering what needs to be accomplished and the benefits and drawbacks of various group process techniques, which interaction methods should be used?
- Is a skilled facilitator or mediator available who is trained in the group process technique and who does not have a vested interest in a particular outcome?
- Is there a plan for circulating materials to members in advance for review?
- Will there be written documentation of the group interaction process?

Recommendations

- Provide appropriate training in group interaction techniques for CDC staff and panelists.
- CDC staff should plan for group interaction activities, assigning activities to groups which are appropriate for group work. Planning staff should avoid encumbering groups with tasks that are more efficiently accomplished by individuals.
- Guideline planning staff should select group interaction techniques appropriate to the specific tasks to be accomplished and the nature of the group or panel. An appropriately chosen group interaction method supports and enhances a group's process, manages bias, and facilitates accomplishment of intended group tasks.
- Circulate materials in advance for review. (For specific techniques such as Delphi, follow exact protocol.) Groups will be more effective if the individual participants have the opportunity to think, perhaps even write, before they are asked to contribute to the work of the group.
- Use skilled facilitators or mediators who are trained in the group process technique selected for the group activity. To add to the credibility of the guideline development process, facilitators and

mediators should not be invested in a particular outcome and should be perceived as neutral and professional by group members. A facilitator with an understanding of the particular topic, but not necessarily an expert, may also serve as the working group coordinator.

- An operational definition of **consensus** should be specified as well as how to present less than full agreement of the panel's findings.
- Maintain good written documentation of the group interaction process, just as good research notes are maintained throughout the research and analytic process. Use group interaction methods which produce written documentation.
- Avoid voting as a method of seeking consensus. A better approach is to rank ideas and to encourage participants to react and add to the work of others. If consensus is appropriate, use mediation techniques.

Approaches to Aiding Group Interaction

Principles

The principles outlined above and elaborated on below provide general suggestions for group planning and process methods rather than support for particular methods. The recommendations reflect a narrative synthesis of the professional opinions of the experts whose work is cited. As the science of group interactive process develops, evidence of the superiority of a particular method applied to guidelines development may become available. Also, new techniques may evolve to enhance group interaction.

- Groups will be more effective if the individual participants have an opportunity to think, and perhaps even write, before they are asked to contribute to the work of the group. Allowing participants to review materials, develop their thoughts, and formulate ideas in advance of a meeting recognizes the diverse styles and needs of group members in how they process and integrate information. Some people work best when they have time to reflect quietly while others are highly stimulated by group discussion. While groups are usually more effective than individuals in generating ideas, individuals are often more effective than groups in developing ideas. Allowing time for individual thought, even during a process such as brainstorming by providing a few silent moments for reflection before discussion, can accommodate the needs of individuals and take advantage of the strengths of group work and individual work.
- Groups generally prefer the opportunity to reflect a range of choices rather than voting. Group members may resist making a fi-

nal selection among ideas, often preferring to offer a range of options rather than a final recommendation. Options generated by participants may need clarifying information or supporting research data that are unavailable at the time of the meeting. Forcing participants to vote to select options may be frustrating when the options cannot be fully developed during the meeting because of lack of information or meeting time limitations. Voting also encourages an atmosphere of competition, stifling the opportunities for participants to react and add to the ideas of others. Instead of voting, it is preferable for participants to rank choices with a technique such as multi-voting i.e, giving individuals multiple votes, where participants rank a set of options. In some instances, it may also be helpful for participants to define their criteria for choosing options, thus identifying important underlying concepts that motivate their individual recommendations.

- The productivity of a group can be enhanced by skilled facilitators. To add to the credibility of the guideline development process, facilitators and mediators should not be invested in a particular outcome and should be perceived as neutral and professional by group members. In instances where politically sensitive issues are at stake or in large meetings such as policy dialogues or public hearings, it may be advisable to use a professional outside facilitator. In other instances, in-house staff trained in group process techniques may serve as facilitators keeping in mind the need for the facilitator to be neutral. A skilled facilitator can structure, organize, and coordinate group meetings to ensure that: (a) all group members are given an opportunity to participate in discussions; (b) power among group members is balanced; (c) the group remains focused on important issues; (d) group meeting time is used effectively; (e) an appropriate written record of the group's work is developed.
- Good written documentation of the group interaction process should be maintained. Just as good research notes are maintained throughout the research and analytic process, written documentation of the group interaction process lends scientific credibility to the process of developing guidelines. It is advisable to use group interaction methods which produce written documentation (e.g., a group memory recorded by a facilitator; participant lists produced by idea writing). Poorly articulated paper trails and chains of evidence can make it difficult to discern from a vague rationale statement which parts of the analytic logic are based on science or theory, the quality of evidence, and how it was interpreted. Written documentation can avoid misleading readers into thinking there is more or less scientific support for guidelines than is actually the case (Pritzker, 1995). Decision analytic models can be used to guide and document decision-making processes.

Few organizations issuing guidelines use formal, structured interactive group techniques to orchestrate the guideline process and make explicit recommendations (US Congress, OTA, 1994). Studies have demonstrated that group composition and aspects of group process become increasingly important as the availability and strength of evidence declines (Lomas, 1988). Many guideline processes are informal and organized around a series of loosely defined steps: (1) a group of experts is assembled; (2) available literature is collected, summarized, and then reviewed by individual panel members; (3) meetings in a “roundtable” format are held where experts express opinions; (4) recommendations are agreed upon, often by a vote; and (5) recommendations are reviewed by outside experts and then are reconsidered by the group. Other approaches to group interaction and decision making are available and two, Nominal Group Technique (NGT) and the Delphi technique, have been studied extensively. Five examples of alternative approaches to group process, including NGT and Delphi, are described below.

Nominal Group Technique.

NGT is a single-question technique. First, a question is formulated and pilot tested to achieve responses with the desired level of specificity and abstraction. Then, group members respond to the question by silently and independently generating ideas in writing. This aspect of the process gives the technique its name—individuals participating in the “nominal” group process are a group “in name only” and do not initially interact verbally. In the next step, a facilitator records individually generated ideas on a flip chart. A serial discussion of the list of ideas follows. Finally, participants rank ideas and discuss the ranking pattern for the group.

This technique works best for groups of 5 to 9 members in situations where there is uncertainty or disagreement about the nature of a problem and possible solutions. The technique was designed to circumvent factors, such as verbal aggressiveness and status, that have an adverse impact on groups (Moore, 1994). The technique has been widely used in human service organizations and in evaluation research (US Congress, OTA, 1994).

Advantages

- Easy to learn and use.
- Gives members a sense of being productive in a short time.
- Generates many ideas quickly.
- Produces a written record simultaneously with the process.

Disadvantages

- Does not allow for in-depth discussion, combination of ideas, or explanation of perspectives unfamiliar to the group. Thus, ideas may not be fully developed in the discussion session.
- Does not focus on a single issue.
- Useful for generating ideas but not for decision making.
- Can polarize a group, encouraging participants to become positional or even adversarial.
- Often “majority rules” and this result may not be optimal in guideline development.

Delphi.

This technique can help groups of experts identify a range of possible program alternatives, explore underlying assumptions or information leading to different judgments, and reach consensus on complex issues. Unlike NGT, the Delphi technique does not require the participants to meet face-to-face.

The Delphi technique is not a method for reaching consensus or making a decision; rather, it is used to inform a decision maker of expert opinion and supporting evidence for consideration (Linstone, 1975). Delphi is useful whenever it is desirable to have pooled judgment. In a conventional Delphi, a small monitor team designs a questionnaire which is sent to a larger respondent group. After the questionnaire is returned, the monitor team analyzes the results and, based upon the results, develops a new questionnaire for the original respondent group. The respondent group is given at least one opportunity to re-evaluate its original answers based upon the analysis of the group response.

A real-time Delphi differs from a conventional Delphi in that, rather than taking weeks to conduct the process, it occurs during the course of a meeting and is often aided by computer terminals for collection and analysis of responses. CDC panels have used a version of real-time Delphi to list, discuss, and rank priorities for issues such as managed care policy.

The Delphi process often varies according to whether the respondent group is anonymous; whether open-ended or closed-ended questions are used in the questionnaires; the number of iterations of questionnaires mailed to respondents; and the decision rules used to aggregate the judgments of the respondent group. The underlying theory of Delphi technique is that improvements in judgment with each Delphi iteration occur because the most knowledgeable panelists confidently retain their judgments and anchor the median close to the true value.

Advantages

- Can be used with a large group.
- Useful when there is a wide range of expertise within a group.
- Members can remain anonymous to each other.
- Produces a written record simultaneously with the process.

Disadvantages

- Not a method for reaching consensus or making a decision.
- May reflect dogma rather than the best judgment of participants.
- Monitor team must be well trained in the technique and neutral in order not to manipulate the process and distort the result.
- Technique may emphasize consensus to such an extent that extreme but useful views are suppressed.
- If conducted by mail or computer, the synergistic effect of face-to-face meetings is lost.
- If a mail questionnaire is used, it takes an estimated 44.5 days to complete a Delphi (Delbecq, 1975).

Idea writing.

This technique recognizes that certain group goals can be achieved best by writing rather than by discussion. Idea writing typically includes four steps: (a) Group organization - a large group is divided into working groups of 4 to 5 individuals; (b) Initial response - participants react in writing to a stimulus question or item and then place their pad (with the initial response) in the center of the group. (c) Written interaction - each participant reacts, in writing, to what is written on each of the other pads. (d) Analysis and reporting - participants read the comments made in reaction to their initial response; the small working groups discuss the principal ideas that emerge from the written interaction; and the group summarizes the discussion in writing. The process is especially useful for large groups and was used successfully with an international planning conference of 700 participants (Moore, 1994).

Advantages

- Controls verbal aggressiveness in a group.
- Allows people the opportunity to clearly phrase their thoughts before speaking.
- Provides immediate written documentation.
- Focuses on a single topic.

- Fast and easy to facilitate.
- Can be done by computer.
- Especially useful for large groups.

Disadvantages

- Participants **MUST** be willing to express themselves in writing (Warfield, 1976; Thissen, 1980).
- Technique does not develop or clarify ideas through discussion.
- Primarily used to react to issues, not for decision making.

Negotiated Mediation Techniques.

The traditional model for rule making is that of agency experts deciding the best way to regulate, offering the public an opportunity to comment, and then issuing binding rules. This process encourages adversarial, uncooperative behavior which was particularly true in regulatory areas affecting the environment and the health and safety of workers. In an effort to improve this process, more government agencies are turning to a consensus-based approach. Variations on this approach continue to evolve (Pritzker, 1995a). Negotiated rule making is done in the context of a FACA chartered committee.

Techniques used in this negotiated rule making process (sometimes called “reg neg”) have potential application in the process of developing guidelines. In negotiated rule making, after an issue is determined to be appropriate for reg-neg, a committee of affected interested parties, including the agency, meets with a skilled mediator. The goal is to reach consensus on the proposed policy. The long-term benefits of the reg-neg process include: more innovative approaches that may reduce compliance costs, earlier implementation, and increased cooperation between the agency and other affected parties. Negotiations that do not result in consensus can also be useful by narrowing issues in dispute, identifying information necessary to resolve disputes, ranking priorities, finding potentially acceptable solutions, and improving the agency’s understanding of the real-world impact of alternative options. Even in programs with no history of adversarial policy making, the agency may obtain a better factual basis for a position and a better understanding of the practical consequences of each option. Negotiations can also help enfranchise parties with important interests.

In government, there is documented evidence that the reg-neg process reduces costs and saves time. Using the reg-neg process, EPA estimates a time saving of 6 to 18 months as compared to the normal rule making process, along with a significant dollar savings from avoiding litigation (Pritzker, 1995b).

The Office of Medical Applications of Research at the National Institutes of Health has used the consensus process in developing clinical guidelines. A 1988 study of a consensus conference on Cesarean birth indicates that a consensus process structured to emphasize scientific evidence lead to panel consensus that reflected this orientation while facilitating decision making (Lomas, 1988). Using a professionally trained mediator to conduct discussions for guidelines panels composed of people with diverse interests aids the communication process and fosters opportunities for full participation by panel members.

Advantages

- Can reduce costs and save time of implementing guidelines.
- Suitable when interested parties hold widely varying interests.
- Consensus-building process that addresses underlying issues as well as achieving a decision.
- Suitable for complex, politically charged decisions that may involve the general public and others.

Disadvantages

- Requires highly trained, professional facilitators.
- May take several meetings over several months.

Decision Analysis, Influence Diagrams, Decision Diagrams.

Decision analysis, influence diagrams, and decision diagrams are types of structural modeling that a group can use to help manage multiple, complex ideas. Decision analytic techniques provide an explicit, quantitative, and systematic approach to decision making under conditions of uncertainty. At CDC, decision models are a component of economic analyses used for policy analysis and guidelines development. Modeling techniques are used to structure policy questions, test hypotheses, illuminate areas where research data is needed, and provide quantitative information of the cost-effectiveness of public health prevention interventions. However, decision analytic methods do not substitute for professional judgment or expertise even though they provide a framework for managing a large array of complex factors.

Groups engaged in guideline development may use decision analytic logic as a guide for reviewing evidence, developing recommendations, and explaining the rationale for guidelines. The methods can be used to define which questions must be answered to arrive at a recommendation, which types of evidence and information are relevant to the analysis, and by what criteria evidence will be judged (Woolf, 1992). (See the section on Analytic Framework.)

Groups may also use decision analytic techniques to structure their process of decision making (Woolf, 1992). In decision analysis, a

graphical representation is created, often in the form of a decision “tree” with “branches” representing different options. Influence diagrams show factors that are important in a decision and diagram the interrelationships between factors in a model resembling a flow chart. Mathematical calculations resulting in estimates of cost-effectiveness or cost-benefit can be performed for models if numerical values are assigned to designate the probability of certain events happening and the utility, or perceived usefulness, of each possible outcome.

Decision models can be helpful to group decision making for several reasons:

- a. **Explicitness.** Unlike intuitive decision making, the decision-analysis process requires that options be clearly stated, consequences be clearly identified, and uncertainties be recognized. Group decision rules may also be made explicit.
- b. **Comprehensiveness.** When making a decision by intuition, a person is only able to simultaneously consider a limited number of options and to process a limited amount of information. When a group of several members is engaged in decision making, the process becomes even more complex. Because the analysis process scrutinizes each part of a decision model for alternatives and outcomes, the process becomes more comprehensive than intuitive decision making can be.
- c. **Improved communication.** To one expert, the word “rarely” may mean a 1% chance; to another, it may mean a 10% chance. The process of decision analysis allows decision makers to understand and convey information clearly about aspects of a problem. With the use of formal decision analysis, it is also easy to document and justify the choices made.
- d. **Facilitated decision making.** It is often difficult for a group to reach decisions, especially complex decisions with far-reaching consequences. The logical, rational process of decision analysis lends structure, organization, and reason to a difficult process.
- e. **Focus is encouraged.** The process of decision analysis encourages group members to focus on truly important issues rather than on issues that merely seem to be important. Structuring a problem can clarify the significant issues involved in a decision. In mathematical models, conducting sensitivity analysis allows the decision maker to consider each variable in the model to determine the importance of the variable in a particular outcome.

For some decisions, elaborate, mathematical models can provide a high degree of structure and organization to the group process of making decisions. In other instances, a group may quickly develop less formal decision models to guide discussions and focus issues during group interaction and decision making.

Advantages

- Can manage a large array of complex factors.
- Provides a high degree of structure and organization.
- Group decision making rules are explicit.
- Comprehensively examines an issue.
- Uses a purely logical base.

Disadvantages

- Elaborate models take much time to develop.
- Serves as a framework but does not necessarily incorporate professional judgment or expertise.

Current guideline development efforts at CDC reflect the variety of approaches to aiding group interaction illustrated by the samples described above. Historically, CDC-sponsored guidelines have been developed by (1) Federally chartered advisory committees (e.g., ACIP), (2) ad hoc expert panels which are convened to develop a particular guideline, (3) staff of the sponsoring CIO, and (4) some combination of the categories described in items 1-3. Group interaction activities have used a variety of formats—from informal discussions to a highly structured Nominal Group Technique process which was used for development of Prevention Effectiveness Guidelines for economic analyses at CDC in 1993.

Table 3 summarizes the characteristics of each of the group interaction methods discussed.

Conclusive evidence suggesting that any of the approaches to aiding group interaction for guideline development is more or less effective than another is not available. For each group interaction method outlined above, evidence supports the effectiveness of the technique in accomplishing specific tasks (e.g., generating ideas for consideration, documenting group interaction, facilitating discussion, reaching a decision). Group interaction methods are tools to facilitate process and not ends in themselves. In addition, the processes are seldom used independently. It is usually necessary to link them together with traditional interacting meetings, document reviews, and other processes. Used correctly, a specific interaction method is selected to fit the group's tasks in the process of developing a guideline. For example, groups occasionally flounder for a period of time over their inability to get started. A successful, productive idea-generating session using a technique such as NGT may help the group perform the task and also contribute to

Current Practice at CDC

Choosing a Group Process Method

Evidence of Effectiveness

Table 3. Characteristics of Group Process Method

	Nominal Group Technique	Delphi	Idea Writing	Negotiated Rule Making	Decision Analysis
Group composition	Suitable for general audience, experts, or mixed	Suitable for general audience, experts, or mixed	Suitable for general audience, experts, or mixed.	A series of techniques and processes. Can include experts as well as general public.	Technical process most suited when there is uncertainty among experts.
Group size	Best for groups of 5-9. Can conduct concurrent small groups.	Best for groups of 5-15. Can be used for groups of 100.	Best for groups of 5-6. Can conduct with concurrent small groups.	Techniques such as mediation are suitable for groups of 2-4. Techniques such as policy dialog or public hearings are suitable for very large groups including the general public.	Typically a workgroup of 2-5 experts constructs the model and analyses results for presentation to policy-maker(s).
Method suitable for achieving consensus?	No. Based on voting.	No. Used to inform a decision maker of expert opinion.	No. Used to generate ideas and document reactions.	Yes. Consensus is the focus of these techniques.	No. Based on decision rules and technical estimates. Informs a decision maker of expert opinion.
Time required for process	1-2 hours for typical meeting.	If done by mail, up to 6 weeks. If aided by computer, 4-8 hours.	1 hour for typical session.	Some techniques such as mediation can take 2 hours to several meetings. The time required depends on the number of parties and the complexity of issues.	Sometimes a quick estimate can be done in a few hours. A full decision analysis may require several months.
Special requirements	None. Easily facilitated.	Monitor team must be trained in the technique and maintain neutrality in formulating iterations of questions.	Very easy to facilitate. Participants must be willing to express themselves in writing.	Highly trained, professional facilitators and mediators required.	Highly trained experts in decision modeling required.
Suitable for in-depth examination of issues?	No. Good for generating ideas, but not for development of ideas or in-depth discussion.	Yes. The use of several iterations of questions allows experts to clarify position, criteria, and recommendations.	No. Good for generating ideas. No development of ideas or discussion. Generally used in conjunction with other meeting techniques.	Yes. The focus on consensus stresses the examination of interests of parties as well as possible options.	Yes. Technique makes explicit all issues to be used as the basis of a decision.

their sense of self-confidence in working together. Selecting group interaction methods which address the various or changing needs of the group can maximize group productivity.

Identifying a Research Agenda

A research agenda is that list of study questions or areas of inquiry that should receive high priority for scientific investigation and funding. A research agenda for public health practice should be established within a broader strategic planning process. The research questions and needs can be identified at several points in the strategic planning process, including: 1) the review of health statistics relevant to the organizational mandate for prevention and control of disease or injury; 2) the review of important outside forces (e.g., social, economic, political, scientific, technological, environmental) to identify opportunities for and threats to achieving the mandate; and 3) the review of organizational capacity to respond to strategic needs. They can also be identified through visioning or scenario-building exercises that identify potential circumstances and allow for a description of the ideal so that steps to achieving it can be more clearly defined (Bryson, 1993).

An evidence-based guideline development process follows the structure of a strategic planning process. An organized approach is used to identify topics based on organizational goals and objectives. An analytic framework identifies the practical alternatives for achieving the goals and provides a description of the causal reasoning linking the proposed practices and relevant health outcomes with evidence of effectiveness (Woolf, 1994). Filling in the data defines the benefits and barriers to each alternative. A spin-off benefit of following an organized approach is the systematic identification of gaps in the scientific foundation for recommendations about practices and their implementation (Woolf, 1994). Explicit frameworks, e.g., decision analysis, lend themselves to identification of knowledge gaps and highlight the importance of specifying uncertainty in estimates and conclusions.

The recognition of inadequacies in data can establish the research needs and be an early step in identifying a research agenda. As one approach to determining research needs, the guideline development process provides a deliberate methodology for addressing pertinent questions and building on existing knowledge of the subject. Gaps in the scientific foundation should be reported in the discussion of the guideline and can provide the basis for a relevant research agenda.

The agenda should be re-evaluated in an ongoing process that continually addresses the most relevant questions. As guidelines are revisited, existing information will be reviewed and synthesized and current gaps in knowledge identified. In this way, a dynamic guideline development process feeds the research priority setting process that will provide the new information necessary to improve the effectiveness of guidelines and renew the cycle.

Example 1

In September 1994 CDC convened a workshop to provide guidance on the public health threat of waterborne cryptosporidiosis (CDC, 1995). Four topic areas were addressed in the workshop and the discussions deliberately focused on summarizing the current knowledge in these areas and defining the information necessary to develop recommendations. Where information was lacking, suggestions were provided for the research methods appropriate to obtain the information. The workshop concluded that current knowledge of waterborne cryptosporidiosis is minimal and does not provide a scientifically sound basis for many of the decisions necessary in the public health response to infection.

Research needs were defined in disease reporting and outbreak study design and in the development of dependable methods for detecting cryptosporidium in drinking water. [Note: while a workshop of experts from a variety of disciplines is not the ideal framework for evidence-based guideline development, it can be useful in situations where the evidence base is limited and timely recommendations are necessary.]

Example 2

Using an evidence-based approach, the U.S. Preventive Services Task Force in 1989 published guidelines for postmenopausal estrogen replacement therapy (USPSTF, 1989). The Task Force concluded that there was sufficient evidence that estrogen therapy can reduce bone loss in postmenopausal women, but that there was insufficient evidence regarding the prevention of fractures, reduction in cardiovascular mortality, and the risk of breast cancer and gallbladder disease. The recommendation was that estrogen therapy should be considered in asymptomatic women at increased risk for osteoporosis and without known contraindications. The 1996 edition of the Guide to Clinical Preventive Services (USPSTF, 1996) revisited the estrogen replacement guideline and updated the analytic framework used to assess this treatment (Figure 2)

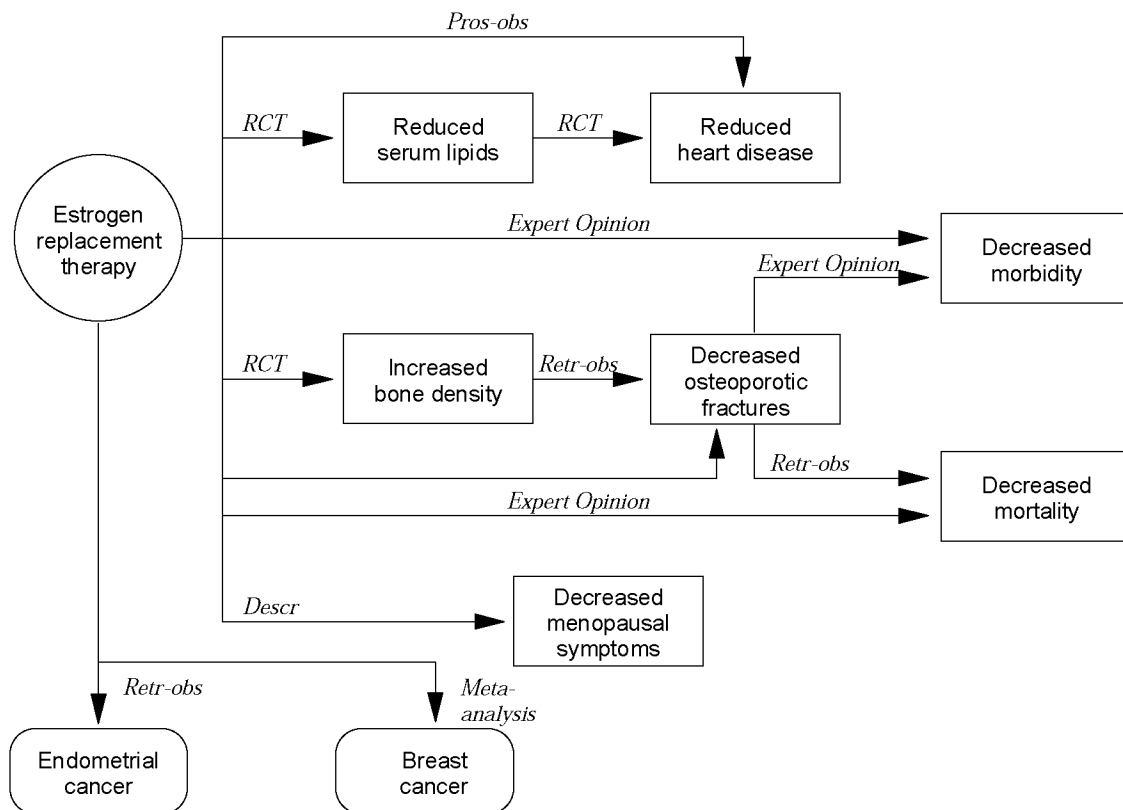
In the new guideline the Task Force concluded that estrogen therapy after menopause relieves vasomotor and urogenital symptoms, produces clinically important improvements in bone density and blood lipids, and is associated with significant reductions in the risk of heart disease and fracture. The Task Force gave a B recommendation (fair evidence to support the recommendation) that clinicians should counsel all women around the time of menopause about the possible benefits and risks of postmenopausal hormone therapy and the availability of treatment options. Additional research needs were identified to include more reliable estimates of the magnitude of the benefits in cardiovascular disease and fracture risk, the appropriate duration of treatment, the benefits and risks in older women (over age 65) and non-white women, the effect of adding progestins, interactions with

other risk factors, and the role of postmenopausal hormone regimens in the risk of breast cancer.

- A systematic guideline development process should identify and help clarify the relative importance of controversies or gaps in knowledge and help to establish a research agenda to determine desired practices and the best implementation strategies.
- The panel should recommend **research** needed to resolve issues of importance, especially those for which full agreement was not reached.
- A mechanism for periodic evaluation of guidelines should exist to facilitate changes in the guidelines, implementation strategies, and the research agenda.

Recommendations

Figure 2. Analytic framework “filled in” with evidence



RCT = randomized controlled trials; Pros-obs = prospective observational studies; Retr-obs = retrospective observational studies; Descr = uncontrolled descriptive studies.

This page intentionally left blank.

Updating Guidelines

How often a guideline should be reviewed and updated depends on several factors: how strongly current scientific evidence supports it, how quickly new scientific evidence is likely to become available, how effective the current guideline is, and changes in the practice environment. A formula for updating guidelines is not appropriate. However, during the development process specific decisions should be made regarding the time frame for updating the guideline. These decisions should consider the strength of existing evidence, the level of controversy surrounding the guideline, and the expectations of new scientific evidence. New guidelines should include a specific time frame for revisiting the guideline for updating or reaffirmation. Unanticipated new evidence may accelerate the review process. Ongoing systematic reviews of the literature including cumulative meta analysis allow “real-time” updating of scientific evidence.

Evidence suggesting that a guideline is not being implemented adequately or that it is not achieving the desired effect should also prompt review and updating. The assumptions on which a guideline is based or the development or implementation process may need re-evaluating. The review should consider the factors discussed in the upcoming section on Encouraging the Adoption of Guidelines. Changes in the practice environment, such as the growth in managed care as a health care delivery system, should also prompt critical review of previously issued guidelines. Relevant perspectives should be addressed to improve interest in and compliance with the guidelines.

The planning of a research agenda also offers an opportunity to revisit guidelines. The systematic assessment of the effectiveness of existing guidelines and of the current scientific evidence can be valuable for identifying pertinent research needs

Answers to the following questions can help determine whether a guideline should be updated.

- How strong is the evidence on which the guideline is based?
- Is new scientific evidence available or expected soon?
- What level of controversy surrounds the guideline?
- Is the guideline effective?

Points to Consider

Recommendations

- Based on the best available information at the time the guideline is issued, a statement should be included that indicates a time frame within which the guideline should be revisited.
- A guideline should be reviewed and updated when new evidence suggests that the guideline is incorrect or is ineffective or strengthens the recommendation.
- Updating guidelines should be considered in the process of establishing a research agenda.

Writing the Guideline

The way guidelines are written can strongly affect their potential for effective use. Planning for a successfully written document must begin early in the process and continue through the cycles of review and revision. The writing process can be more easily managed if a work plan and time frame are developed and followed. Keeping to a schedule is important—getting timely reviews by setting a clear and reasonable time frame for review, evaluating comments, and incorporating the comments into the document. Good planning and adherence to the schedule can improve the final product.

In writing the guideline, a compelling case needs to be made to the users of the guideline to effect its implementation. To do so, the document must answer the potential users' questions. Major questions that all relevant users (e.g., society, clinicians, patients, health care plans) should be able to answer include the following: Is the problem relevant to me? Is it an important problem? Can I do something about it? Is it worth doing? Decisions or opinions that are based on weak evidence or expert opinion should be duly noted.

Guidelines should provide clear recommendations and employ credible methods that document the rationale on which the recommendations are based. The guidelines should be practical, yet specific, comprehensive, and flexible enough to be useful in everyday practice. The language, logic, and symbols used in the guidelines should be unambiguous and easy to follow. The length should be appropriate to the subject matter and the intended audience.

Formatting is defined as the presentation of guidelines in physical arrangements or media that can be readily understood and applied by the intended users of the guideline. Effective formatting means the guidelines will be delivered to the intended audiences in a way that promotes the reception, understanding, acceptance, application, and a positive impact. Guidelines can vary quite dramatically, both logically and graphically, in their modes of presentation. The major approaches are freetext and formalized presentations, including if/then statements, algorithms, flowcharts, and decision trees (Field and Lohr, 1992).

No matter which format is used, all public health practice guidelines should contain basic components in their core presentation. These components include the relevance of the health problem, the magnitude of the problem, the nature of the intervention, the guideline development methods, the strength of the evidence, cost effectiveness, a discussion of implementation issues (including current extent of implementation), evaluation, and the recommendations of others.

An explanation of how the guideline was developed and who participated in the process should be provided (Field and Lohr, 1992) along with a contact person. Also, a one-page summary should be provided that addresses all of the basic components. The summary should be written using lay terminology, and it should be fully accessible to the general public. These components are considered the minimal requirements of an effectively written guideline; they are not meant to be comprehensive. (A discussion of these basic components follows the Points to Consider and the Recommendations.)

Points to Consider (Preparing to Write)

Several questions should be answered during the writing of a guideline.

- For whom is the guideline relevant? (For example, are the guidelines for physicians, the public, healthcare workers, a combination audience, etc.?) Knowing the intended audience will play a significant role in how the guidelines are written, the approach, language, and style. In the case of public health practice guidelines, the audience will be public health practitioners, such as state and local health department program managers and policymakers.
- Has the magnitude of the health problem been conveyed adequately?
- Has the guideline development process been described in sufficient detail to establish the credibility of the methods and the results of the data synthesis?
- Has the case been made to the intended audience that the quality improvement in public health practice is worth the costs of implementing the guideline?
- Have implementation strategies been identified that establish the feasibility of implementing the guideline?
- What measures and vehicles are available to monitor the extent of implementation and the effectiveness of the guideline?
- Have the recommendations of other relevant organizations been addressed?
- How will the guidelines be communicated to potential users? The main dissemination vehicles for CDC public health practice guidelines are the MMWR/RR, journal articles, and the Federal Register. In addition, press releases on the guideline content are issued at the same time.

Recommendations (Author Responsibilities)

The following recommendations apply to all categories of guidelines regardless of the guideline development process.

- Writers and editors should be familiar with the subject matter or have the ability to grasp the information, interpret it, and explain it (medical/science writers/editors).
- Length, format, layout, and style should be suitably matched to the intended audience, subject matter, intention of the guideline, and dissemination vehicle.
- A work plan with time frame must be thoroughly thought out and adhered to so as to keep the project on schedule.
- CDC public health practice guidelines should address certain basic components: the relevance of the health problem, the magnitude of the problem, the nature of the intervention, the guideline development methods, the strength of the evidence, the cost effectiveness, a discussion of implementation issues, evaluation issues, and the recommendations of others. A brief but complete summary and a contact source should be provided.

Since the main audience for public health practice guidelines is public health practitioners, such as state and local health department program managers and policymakers, and the main dissemination vehicles are the MMWR/RR, journal articles, and the Federal Register, the core guideline should be written at the level of professional practitioners in public health and clinical medicine. The document can subsequently act as the source for different presentation vehicles (brochures, news release, and training materials) for specific audiences. In adapting the original document for different audiences, changes may need to be made, depending on the audience and presentation planned. Further discussion of the core components of guidelines follows.

The guideline should state explicitly for whom the recommended practices apply. A guideline is relevant to those who carry out the recommended practices (clinicians, sanitarians, health educators, etc.), those who benefit from the practice changes (i.e., the population affected), and those who assume risk or financial cost for the problem (health care purchasers, employers, local governments, etc.). A checklist (Table 4) should be completed to denote relevant categories for crossreferencing of the guideline.

The magnitude of a health problem can be defined in several ways and the greater the number of perspectives addressed, the broader will be the recognition of importance. Health problems that affect large numbers of people, have catastrophic consequences, or require costly control interventions are generally recognized as important. Other factors that influence the perception of importance include the transmissibility of the health event and potential for outbreaks, perceptions of blame (whether the responsibility of an individual (e.g., a behavior) or an effect beyond individual control (e.g., an environmental hazard)), and the dramatic appeal and media attention. The importance of the

Basic Components of a Written Guideline

Relevance of the Health Problem

Magnitude of the Problem

health problem should also be considered in terms relevant to the persons or organizations for whom practice changes are being advocated. The magnitude of the health problem can be reflected through measures of absolute risk, relative risk, severity, transmissibility, popularity, and costs.

Nature of the Intervention

A clear and explicit description of what practices are being recommended should be provided in a distinctively marked location and preferably early on in the written document. The objectives of the guideline should be included in this description. By presenting the guideline before the supporting documentation, a context is provided for assessing the documentation.

Guideline Development Methods

The methods used to arrive at the strategy being presented as a guideline should be described in detail. Included should be a description of the decision-making body, their affiliations and their potential conflicts of interest; the methods used to select the decision-making body; the analytic framework used for choosing between guideline options; and the methods used for synthesis of existing data. The technical and other reviewers of the guideline document should be identified.

Effectiveness of the Intervention

What are the reasons or principles on which the recommended course of action is being based? What can be expected from implementing the suggested recommendations? A clear explanation of what the scientific evidence shows about the effectiveness of the intervention being proposed by the guideline should be provided. The results of the data synthesis should be provided as justification for the decisions made. Recommendations should include explicit statements about the strength of each recommendation and the quality of the supporting evidence (see page 56 for example from the USPSTF). The data synthesis should address such issues as feasibility of early recognition, efficacy of the intervention, effectiveness of interventions (depends on adherence rate, scale of adoption, and utilization rate in general population), approaches to implementation of the guideline and the ability to generalize those approaches, and conclusions regarding the strength of the evidence supporting the guideline. A flow chart of the practice options considered in the decision process would be worthwhile.

Cost Effectiveness

The guideline should include a statement on the cost effectiveness of implementing the proposed practices from the societal perspective. While the societal perspective is important for establishing public policy, many individual practitioners and consumers for whom the guideline is intended will not make decisions based on the societal implications and other appropriate perspectives (e.g., health departments, health plans, employers, and consumers) should also be addressed. When data are lacking to address a relevant perspective, the absence of data should be noted and the process used and conclusions drawn regarding value should still be presented. In establishing the value of a guideline, other considerations include opportunity costs and downstream or domino effects, consumer satisfaction and quality

of life, health services quality indicators (potentially used to compare providers), and liability implications.

The interventions being proposed should be explained in the context of public health practice experience. Discussion of implementation issues should facilitate user decisions about the feasibility of adopting the guideline. Who will need to do which work to implement the guideline? How can the guideline be integrated into current practices without substantial changes in staffing or routines? Is there flexibility for clinical judgment? What is the current level of implementation? What strategies could increase compliance with recommendations? What resources and procedures are necessary to maintain compliance?

How can successful implementation be measured? What level of compliance is realistic? Considerations for performance measurement should be addressed, including the tests and procedures for monitoring the implementation and the effect of the guideline. Describing the methods which identified the problem may be useful for defining the measures important for performance measurement (measures for surveillance).

If other groups have recommendations in this area they should be mentioned and discussed in the context of how they conform or differ from the recommendation being proposed. If other groups are endorsing this guideline that should also be mentioned.

A contact for handling inquiries and requests should be provided at the end of the document.

Implementation

Surveillance and Evaluation

Recommendations of Others

Agency Contact (for more information)

Table 4. Guideline Cross-Referencing Checklist

<u>Setting</u>	<u>User</u>	<u>Target Population</u>
<input type="checkbox"/> clinic <input type="checkbox"/> health department <input type="checkbox"/> medical office <input type="checkbox"/> laboratory <input type="checkbox"/> hospital <input type="checkbox"/> workplace <input type="checkbox"/> community <input type="checkbox"/> managed care plan <input type="checkbox"/> health agency <input type="checkbox"/> home <input type="checkbox"/> school <input type="checkbox"/> legislature	<input type="checkbox"/> physician <input type="checkbox"/> nurse <input type="checkbox"/> other health care provider <input type="checkbox"/> health official <input type="checkbox"/> sanitarian <input type="checkbox"/> laboratory worker <input type="checkbox"/> researcher <input type="checkbox"/> health educator <input type="checkbox"/> advocate/ community-based organization <input type="checkbox"/> legislator <input type="checkbox"/> consumer/patient	<input type="checkbox"/> children <input type="checkbox"/> adolescents <input type="checkbox"/> elderly <input type="checkbox"/> women <input type="checkbox"/> pregnant women <input type="checkbox"/> minorities <input type="checkbox"/> persons with disabilities <input type="checkbox"/> inner-city residents <input type="checkbox"/> workers (farm, etc)
<u>Topic</u>	<u>Intervention</u>	
<input type="checkbox"/> infectious disease <input type="checkbox"/> chronic disease <input type="checkbox"/> environmental health <input type="checkbox"/> injury <input type="checkbox"/> occupational health <input type="checkbox"/> behavioral risk factors <input type="checkbox"/> reproductive health <input type="checkbox"/> health planning <input type="checkbox"/> international health	<input type="checkbox"/> prevention <input type="checkbox"/> screening/diagnosis <input type="checkbox"/> treatment	

Obtaining Critical Reviews and Public Comments

Obtaining critical reviews and public comments are essential steps in devising guidelines that are credible and useful and in anticipating the challenges of acceptance and implementation (Field and Lohr, 1992). A multidisciplinary review process should include all key groups from the scientific, advocacy, and provider constituencies. Although potentially requiring more time and funds, such an approach can also help developers of guidelines better understand the situations in which guidelines will be implemented (Field and Lohr, 1992).

Review of draft guidelines by experts in the field and relevant organizations and agencies that are involved in the area, or will be involved in the implementation, can provide broad input on content and policy issues. Review by relevant content experts can ensure scientific and clinical validity (Woolf, 1992). In addition, CDC's Office of the General Counsel (OGC) usually reviews guidelines for legal considerations, references to liability or litigation concerns, discussions of federally chartered advisory committees, references to laws, and other issues related to the functions of the office.

In selecting groups for review and comment, many aspects need to be considered. They include the legal and social implications of the guideline, the scope of the guideline and all the groups and agencies that may be affected, both directly and indirectly. Questions that need to be answered include whether political support or regulatory follow-up will be needed and what role the advocacy community may play in the acceptance of the guideline. Activities such as draft document circulation, draft publication, and hearings can provide opportunities for including groups in the review and comment process. In some cases pretesting may be beneficial. Pretesting involves asking a small sample of practitioners to use a guideline for a brief period and then collecting their suggestions on ways to improve the document.

Identification of reviewers for draft guidelines is an area in which groups may differ substantially in how they select participants, assuming that they have a process for reviewing draft documents at all. But whatever the approach, the process serves to broaden input (Field and Lohr, 1992).

In addition to review by experts in the field, proposed guidelines are often published in the Federal Register (FR) with a period available for comment. Guidelines publication in the FR is not usually required by law; it is used, however, as a convenient means for broad dissemina-

tion. Open meetings or hearings are often scheduled for public comment.

Some guideline-related reports have specific legal requirements regarding review and comment. ATSDR, for example, is required by the Superfund Amendments and Reauthorization Act to have all studies and results of research, other than health assessments, peer reviewed before they are reported or adopted. The peer review for ATSDR must be completed, to the maximum extent practicable, within 60 days. In the case of research conducted under the National Toxicology Program, such peer review may be conducted by the Board of Scientific Counselors. In the case of other research, such peer review must be conducted by panels consisting of no less than three nor more than seven members who are disinterested scientific experts selected for such purpose by the administrator of ATSDR or the administrator of EPA, as appropriate, on the basis of their reputation for scientific objectivity and the lack of institutional ties with any person involved in the conduct of the study or research under review. Support services for such panels must be provided by ATSDR or the EPA, as appropriate.

Just as important as knowing who to involve in the review process is knowing when to involve them—before, during, or after the expert panel has completed its deliberations.

Points to Consider

- Will scientific review by content experts and policy review by affected stakeholders be pursued simultaneously?
- Has a formal plan been developed for obtaining reviews by content experts?
- Is public comment needed? If so, has a plan been developed for obtaining it?
- Have all the stakeholders in the process been identified and has consideration been given to the optimal time to bring them into the process?
- Should the guideline be pretested and revised during the development process?
- Are there specific legal requirements that need to be met?

Recommendations

The following recommendations apply to all categories of guidelines regardless of the decision-making authority.

- Use of experts to review draft guidelines is especially critical in anticipating the challenges of implementation and in devising guidelines that are credible and useful.

- The involvement of a multidisciplinary group representing all stakeholders can help encourage acceptance of the guidelines.
- Draft guidelines should be reviewed by relevant content experts to ensure epidemiological, statistical and clinical validity as well as by relevant organizations and agencies to provide broad input on content and policy issues.
- An effective and responsive guideline review and comment process should include public input, if appropriate.

Public and private agencies which sponsor the development of guidelines have used several approaches for obtaining critical reviews and public comment. Open meetings, publication of the proposed guideline for comment, peer review, and solicitation of input from specific individuals or groups are examples of the most common approaches.

Historically, CDC-sponsored guidelines have been developed by (1) Federally chartered advisory committees (e.g. ACIP), (2) ad hoc groups of individual consultants which are convened to develop a particular guideline, (3) staff of the sponsoring CIO, and (4) some combination of the categories of decision makers described in items 1-3.

Examples of how the groups listed in items 1-3 above have sought review and comment are described below.

The ACIP is an example of a Federally chartered advisory committee. Committee membership includes ex-officio members from the NIH, FDA, and liaison members from most organizations which are subject to or will implement the recommendations. All recommendations are developed only after extensive information is gathered (research data, surveillance reports). Working groups are often formed to extensively analyze the research data for presentation to the full Committee. ACIP's process for obtaining critical input includes extensive review by staff of the CDC, ACIP members and liaison members, and outside expert consultants at all stages of guideline development. Public comments are solicited during the open committee meetings and are considered in the decision-making process.

The Working Group which prepared the USPHS and Infectious Disease Society of America (IDSA) Guidelines for the Prevention of Opportunistic Infections in Persons Infected with Human Immunodeficiency Virus is an example of an ad hoc panel developing guidelines. This group's process involved having the draft recommendations reviewed by consultants from CDC, NIH, and IDSA, as well as by members of other Federal and non-Federal agencies, community organizations, physicians caring for HIV-infected persons, and HIV-infected persons themselves. These recommendations were discussed at a two-day meeting convened by CDC, NIH, and IDSA and comments were solicited from the public. Final recommendations were ap-

Approaches to Obtaining Critical Reviews

proved by USPHS and IDSA. The recommendations were also reviewed and endorsed by the American Academy of Pediatrics, the Infectious Diseases Society of Obstetrics and Gynecology), and the Society of Health Care Epidemiologists of America.

The Guidelines for School Health Programs to Prevent Tobacco Use and Addiction is an example of guidelines developed by CDC staff. These guidelines were developed by CDC staff in collaboration with experts from 29 national, federal, and voluntary agencies and with other leading authorities in the field of tobacco-use programs. To develop these guidelines, CDC staff convened meetings of independent expert consultants from the fields of tobacco-use prevention and education, reviewed published research, and considered the conclusions of the National Cancer Institute Expert Advisory Panel on School-Based Smoking Prevention Programs and the findings of the 1994 Surgeon General's Report, "Preventing Tobacco Use Among Young People."

Encouraging the Adoption of Guidelines

A potential user adopts a guideline by deciding to accept and implement the recommended practices. Guideline developers can influence a user's decision to adopt the recommended practices by ensuring that the guideline is appealing and widely disseminated (Figure 3). A well-conceived guideline should anticipate the needs and concerns of potential users and build a compelling case for the recommendations (Field and Lohr, 1992). There are many factors, in addition to scientific evidence, that influence the appeal of guidelines (US Congress, OTA, 1994). These factors include the relevance of the health problem, the importance of the health problem, the credibility of the process used to derive the recommended practices, the feasibility of implementing the guideline and the value of implementing the guideline. Dissemination strategies include education through publication in journals, the lay news media, the electronic media, continuing education as with clinical peer detailing, and consumer information and incentives, both administrative and financial (US Congress, OTA, 1994).

As acknowledged in the Institute of Medicine's study on guidelines for clinical practice (Field and Lohr, 1992), scientific knowledge will continue to grow but will always remain an incomplete foundation for guidelines and implementation strategies. Although neither comprehensive nor prioritized, the issues and methods outlined in this section can affect the adoption of guidelines.

- To get the attention of those who need to review and apply guidelines, the users and the populations for whom the guidelines apply must be clear.
- The health importance should be described from the range of relevant perspectives.
- Confidence in the guideline should be established through a description of the methods used, a summary of the existing scientific evidence, participation and endorsement of credible individuals and institutions, and a description of relevant recommendations and guidelines already existing.
- The purpose and presentation of the guideline should be clear.

Issues and Methods

Relevance:

Importance:

Credibility:

- Feasibility:**
- When possible, guidelines should address a single, well-defined practice.
 - Guidelines should attempt to increase the performance of all users, not just the worst performers.
 - Guidelines should leave room for practitioner judgment in a manner that is consistent with the strength of the scientific evidence and with variation in risk and effectiveness in different populations.
- Value:**
- The costs and benefits of implementing a guideline should be presented from the societal perspective and other relevant perspectives, (public health clinics, local government, managed care organizations, private practitioners).
 - Cost-effectiveness data should be used, when available, to influence purchasers and providers of care to establish financial and administrative incentives for use of the guideline.
 - The legal implications of a guideline should be considered in its development and dissemination.
- Dissemination:**
- Active dissemination strategies should be used for guidelines, when possible, including education of providers through local information networks and respected colleagues, consumer information campaigns, and administrative mechanisms to facilitate practice changes.
 - The electronic Prevention Guidelines Data Base should be kept current as a venue for guideline dissemination.

Points to Consider

Appeal of the Guideline

Relevance

A guideline should include a clear statement of its audience. The use of a checklist of categories (Table 4) can facilitate determination of relevance and cross-referencing of the guideline.

Importance

It is incumbent on those preparing guidelines to assure that the pertinent information establishing importance is readily accessible to the users. Parameters reflecting the importance of the health problem include:

- absolute risk: the total number of cases, incidence, and prevalence
- relative risk: the population subgroups disproportionately affected

- severity: magnitude of the event, case-fatality ratio, drama of case reports, prematurity of outcomes, impact on quality of life
- transmissibility: propensity to cluster, to strike without warning, to be affected by environmental factors beyond individual control
- public interest/concern: which may be due to celebrity involvement or media attention
- costs: health care and human/social costs from all relevant perspectives (e.g., society, geopolitical unit, health system or office, employer, individual)

Magnitude of Effect

A guideline that promises a major improvement in health or prevention of a catastrophic outcome will be more readily accepted than one that promises only a marginal effect (US Congress, OTA, 1994). The magnitude of effect depends on both the health importance of the problem and the effectiveness, not just efficacy, of the intervention. A guideline should state the magnitude of the effect explicitly and should address as many different provider and consumer perspectives as possible on importance.

Aspirin prophylaxis against myocardial infarction is an example of importance and effect making for an appealing recommendation. The incidence, morbidity, and mortality of coronary artery disease are substantial. Adoption of aspirin prophylaxis for those at risk of coronary artery disease was considerable (Krumholz, 1996) following a study of U.S. male physicians that found a 47% reduction in the risk of fatal and nonfatal myocardial infarction (Steering Committee of the Physician's Health Study Research Group). The importance was readily apparent from the perspective of public health, health care providers, and patients. The effectiveness of this intervention was enhanced by its simplicity and low cost, despite a possible increased risk of hemorrhagic stroke.

Credibility of the Guideline

Credibility comes from the quality of the methods used to develop the guideline, the credibility of the participants and sponsors, the strength of the scientific evidence supporting the guideline, the level of agreement among existing guidelines and opinions, and the clarity of presentation.

Quality Development Process

The methods used to develop the guidelines should be readily apparent and clearly stated. The methods include the criteria used for selecting the decision-making body, the analytic framework used for identifying and choosing between guideline options, the approach to synthesis of existing data, and measures implemented to minimize real

and perceived conflicts of interest. The more systematic and demanding the methods, the more appealing the guideline is likely to be. If these methods are used in the development of serial guidelines (e.g., for a Federally Chartered Advisory Committee such as the ACIP), the methods should reference a readily accessible location where they are described in detail.

Respected Collaborators and Supporters

Guidelines issued by respected sources are more appealing than those from less credible sources. Credibility may be increased by broadening the constituency of organizations and individuals endorsing and promoting the guideline (Field and Lohr, 1992; US Congress, OTA, 1994; Weingarten, 1995). Medical associations (e.g., the American Academy of Pediatrics) frequently participate and endorse CDC guidelines. Depending on the intended users, other participants may be viewed as credible sources. If managed care is considered an important constituency, the endorsement of large health maintenance organizations or their trade organizations may be worthwhile. Similarly, consumer organizations and business groups could be valuable allies in promoting a guideline and endorsement by credible consumer or business organizations may enhance the appeal of a guideline. Multi-disciplinary participation can involve hearings, review, pilot testing, and other activities that support the guideline development and dissemination process (Field and Lohr, 1992).

Strength of Evidence

Decisions to change practices are usually made at the local or user level (Field and Lohr, 1992; US Congress, OTA, 1994). The more compelling the scientific evidence that supports the effectiveness of a guideline and an implementation strategy, the more likely the guideline will be accepted as a standard of care. This principle applies both to individual providers and to systems of care that establish their own practice guidelines, such as managed care plans. Strong scientific evidence reduces the opportunity for conflicting recommendations coming from different interest groups with varying perspectives. When scientific evidence is lacking or equivocal, the appeal of a guideline depends on non-scientific issues and the credibility of the proposing body is vulnerable. Guidelines should summarize the quantity and quality of scientific data and make clear the role of these data in their development (Field and Lohr, 1992).

Guideline Agreement

A guideline that reviews relevant recommendations and guidelines of other groups affords a context for its interpretation (Field and Lohr, 1992). In this context, differences can be explained and justified. If a guideline differs from prevailing opinion or existing guidelines, then the strength of other factors such as the scientific evidence, must be particularly strong. Addressing competing guidelines can add credibil-

ity and can highlight the strength or weakness of current scientific evidence (US Congress, OTA, 1994).

Clarity

Guidelines will be more effective if they are explicit, and concisely worded. (US Congress, OTA, 1994) Clarity can be improved by format as well as vocabulary and syntax. A format with easily identified subsections that address pertinent components, such as the characteristics described in this section, can improve the comprehension and usefulness of the guideline.

Feasibility of Implementation

Feasibility reflects the factors that make implementation easier. Such factors include keeping recommended practices as simple and focused as possible, integrating them into current practice as much as possible, accommodating effective quality improvement principles, addressing exceptions to recommended practices and flexibility for judgment, and describing potential administrative mechanisms to enhance implementation (US Congress, OTA, 1994).

Simplicity

Guidelines that attempt to change a single, well-defined practice tend to be more successfully incorporated into practice than guidelines that address multiple behaviors and vaguely defined practices (US Congress, OTA, 1994). Unfocused guidelines that attempt to achieve multiple, diverse objectives can become complex and lack clarity of purpose. For example, a guideline that recommends annual or biennial mammography for all women between 50 and 75 years of age will more likely be implemented than a strategy calling for annual breast examination beyond age 40 with patient counseling about breast cancer and teaching of breast self-examination. Differences in effectiveness aside, the latter recommendation includes multiple, complex activities that are poorly standardized.

Simplicity can also be addressed in the degree to which changes in current practice are called for in the guideline. A guideline that is integrated into current practices without substantial changes in staffing or routines will be more appealing than one requiring major structural changes (US Congress, OTA, 1994). Similarly, practice changes that require minimal effort to maintain will be more appealing than those requiring a high level of diligence and resources on an ongoing basis.

Quality Improvement

The opportunity to improve the quality of care is a fundamental incentive for implementation of a public health guideline. When possible, guidelines should suggest methods for incorporating quality monitoring and improvement. For example, a key principle of continuous

quality improvement (CQI) is that guidelines are most effective when they focus on system changes designed to raise the average performance of all practitioners, as opposed to isolating and penalizing the worst performers (US Congress, OTA, 1994). Missed opportunities for vaccination could be tracked and physicians with low rates could be given incentives for increasing vaccinations. Alternatively, the reasons for under-vaccination could be determined and changes recommended that would improve the ability of all physicians to vaccinate patients in need. If one of the problems is not knowing who needs a vaccine at the time of a visit, a flagging system could be proposed that would improve identification of children due for a vaccine so that it can be ordered (US Congress, OTA, 1994). This second approach gets at the root of the problem and applies to all providers and not just the “bad performers.”

The ability to track performance is also central to CQI. The success of CQI relates to data-oriented decision making and ongoing feedback so that necessary modifications in the guideline can be implemented (Field and Lohr, 1992). Implementation of guidelines at both the institutional level and the national level should be considered part of an iterative process that includes ongoing inquiry, refinement, and implementation. Proposed guidelines should include realistic methods for measuring the implementation and effectiveness of guidelines in order to facilitate this improvement process.

Flexibility

In the absence of conclusive scientific evidence, guidelines represent current best practices, irrespective of the rigor of the development process. A guideline should leave room for clinical judgment by practitioners in a manner that is consistent with the strength of the scientific evidence and the variation in individual risk and in practice effectiveness in different situations. The weaker the evidence and the greater the variation, the greater should be the flexibility. A guideline should address possible exceptions to recommended practices and the role of preferences of relevant parties (patients, communities, employers, etc.) (Field and Lohr, 1992; US Congress, OTA, 1994).

Administrative mechanisms

Purposeful structuring of administrative rules and procedures can facilitate practice changes (US Congress, OTA, 1994). Delineation of potential administrative mechanisms within guidelines and recommendations may improve their implementation. Drug formulary committees can influence the proper use of medications by restricting access to certain drugs (US Congress, OTA, 1994). Automated reminders on computers and in medical records can increase awareness of need for regularly provided services, like immunizations and cancer screening (US Congress, OTA, 1994). Standardized forms can also serve as reminders for proper care in specific situations, such as a checklist for procedures and counseling during prenatal care visits or prescription

information for certain medications. Administrative policies can increase guideline compliance by making procedures easier. Influenza vaccination was increased in settings where nurses had explicit instructions for identifying and counseling patients and standing orders to provide influenza vaccination without active ordering by physicians (US Congress, OTA, 1994). Administrative and economic approaches may be particularly effective when desired practices are clear and supported by consistent scientific evidence. A guideline that identifies a variety of administrative mechanisms can facilitate implementation by a wider range of users.

Value of the proposed intervention

Value incorporates cost with the attributes of relevance, importance, quality, and feasibility. Value describes the net benefits of the proposed practices per unit of cost. A guideline should include a projection of the value of implementing the proposed practices from a variety of perspectives (Field and Lohr, 1992). This statement of value should help potential users clarify their personal or organizational reasons for implementing the guideline.

Perspective

The perspective defines the costs and benefits to be considered in implementing a guideline. A societal perspective is the broadest and considers costs and benefits from the standpoint of the entire population (i.e., society). Public health and CDC would traditionally consider the societal perspective most appropriate for their constituency. Nonetheless, attempts to intervene are usually at a level much smaller than the entire society, such as state or local health departments or hospitals, health maintenance organizations, and individual physician offices. While the societal perspective may be appropriate for the ultimate purpose, the perspective of the health units being enlisted to achieve that purpose should also be considered (Field and Lohr, 1992). When the societal perspective drives the final decisions for the guideline, that decision should be explained and the impact of the guideline should be estimated from the other relevant perspectives. Altruism and public relations may drive some decisions of large health delivery systems, but ultimately, their interests are in the costs and benefits to their institutions and to their patients. Many clinical practitioners will not appreciate the national public health perspective if the implications for their practices and individual patients are not clear. The easier the costs and benefits can be identified from the users' standpoint, the more readily the guidelines will be considered and implemented.

Legal Implications

Public health practice guidelines published by the federal government are frequently cited as evidence of the legal standard of care in litigation settings. As such, recognition by a court that a federal recommendation is part of the standard of care in a given community provides a

significant incentive for implementation. Standards of care carry liability implications that are powerful incentives for implementation (US Congress, OTA, 1994; Weingarten, 1995). Legal considerations can be a two-edged sword, however, as guidelines that are controversial can be contested in the judicial or litigation arena where scientific evidence may not be the primary concern (e.g., motorcycle helmet laws). Anticipating the possible legal implications of guidelines is important. Accordingly, OGC should be consulted throughout the process of developing guidelines. However, legal considerations should always be secondary to medical and scientific evidence relevant to the proposed guidelines.

Dissemination Strategies

Passive Dissemination

Passive dissemination of research findings and recommendations occurs through publication in professional literature, the lay media, and more recently, through electronic media such as the Internet. Current CDC guidelines are now accessible by the public through the Prevention Guidelines Data Base within the CDC home page on the World Wide Web. Passive dissemination is not, by itself, an effective means of encouraging implementation unless the findings are dramatic and the effect is widely agreed to be clinically important and readily applicable. The use of aspirin for prophylaxis against morbidity and mortality from myocardial infarction is a rare example of effective passive dissemination. Myocardial infarction is widely accepted to be an important condition and the intervention of aspirin taken prophylactically proved to be simple, inexpensive, and highly effective. In this instance, the scientific evidence and the source were credible and the findings were broadly publicized in the lay press, raising the awareness of consumers as well as providers. These circumstances led to the substantial adoption of recommendations for aspirin prophylaxis in persons at high risk of myocardial infarction (Krumholz, 1996). Generally, however, more active strategies will be necessary to improve adoption of guidelines (US Congress, OTA, 1994).

Clinical Peer Detailing

Local information networks and the recommendations of respected colleagues are an important influence on clinical practice because nearly all clinical decisions are multi-factorial and depend on the unique characteristics of the patient population and the practice environment (US Congress, OTA, 1994). Gaining endorsement from respected colleagues can allay some reluctance to change (Weingarten, 1995). A specific approach to dissemination via peers is “detailing.” Detailing combines attractive dissemination features of one-on-one interaction with a credible source and message tailored to the needs of the provider. Educational interventions are most likely to be successful in increasing guideline implementation when they provide opportu-

nity for personal interaction, include respected clinical leaders, identify specific learning objectives, and when a guideline has strong support among colleagues (US Congress, OTA, 1994).

Consumer Information

In addition to changing the knowledge, beliefs, and practices of providers, consumer education is a valuable strategy for guideline dissemination. When those who would benefit from a guideline are informed and demand the service, they are likely to influence provider decisions (US Congress, OTA, 1994). Health care systems, particularly managed care, are increasingly addressing patient satisfaction, and information can increase the patient's sense of control and satisfaction with care. Not only can consumers be valuable advocates for guidelines, but the information strategy serves an important ethical obligation for informed consent (US Congress, OTA, 1994).

Financial and Administrative Incentives

Financial incentives for adherence to a guideline appear to be effective at both the individual provider and the institutional level. Favorable reimbursement policies can motivate changes in practice by providers (US Congress, OTA, 1994). The reduction in length of hospitalization following implementation of prospective, diagnosis-related payment for inpatient services under Medicare in 1984 demonstrates how financial incentives at the institutional level can effectively change provider practices (US Congress, OTA, 1994). Performance report cards, such as the Health Plan Employer Data and Information Set (HEDIS), are indirect incentives used by purchasers of health care to assure adherence to quality care standards by health plans (National Committee for Quality Assurance, 1996). Competitive and reimbursement incentives such as these at the organizational level can be translated to administrative and financial incentives at the provider level.

Practice profiling, the data-oriented review and feedback of provider practice patterns and outcomes, is a specific administrative mechanism designed to reduce practice variation and bring conformity with guidelines and standards of care (US Congress, OTA, 1994). The strategy is most effective when individualized, presented by a respected peer, and provided in a context for peer practices. The providers must agree with the proposed practices. This approach tends to be more effective when attempting to increase a particular practice, in contrast to discontinuing an unnecessary practice (US Congress, OTA, 1994). The feedback information is most useful when it identifies a process that can be changed (e.g., below average use of inhaled steroids for asthmatics) instead of an outcome that may have numerous interrelated causes (e.g., asthma death rates).

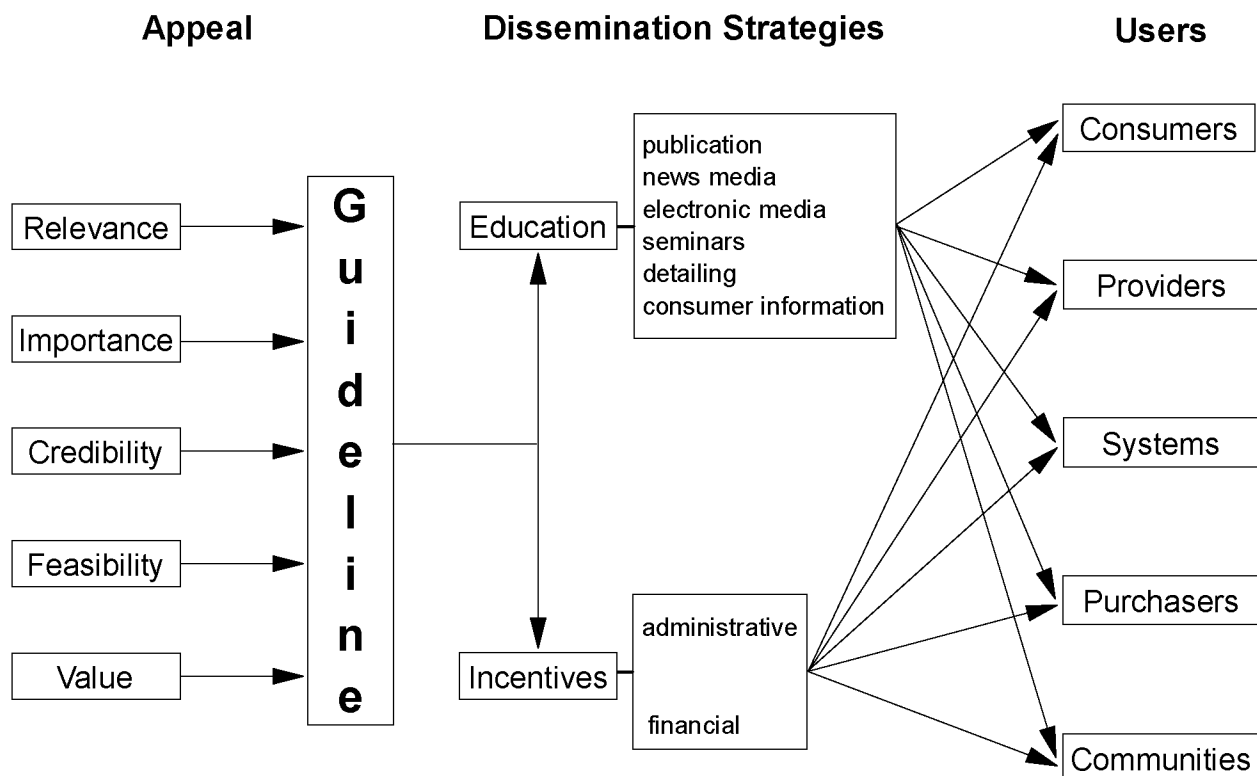
Intensity of Dissemination

The greater the dissemination effort, both in terms of venues and quantity, the more effective will be the implementation of a guideline (US Congress, OTA, 1994). To be successful, a guideline must address the range of influences relevant to the practice being addressed. Combinations of intervention strategies are more effective than single interventions strategies.

- Follow-up of Dissemination

Continued efforts must be made to assess the effectiveness of dissemination efforts and to identify opportunities for changes (US Congress, OTA, 1994). Even effective dissemination strategies may benefit from ongoing promotional efforts to remind users of the guideline.

Figure 3. Framework for Encouraging the Adoption of Guidelines



References

- Advisory Group: Council on Linkages between Academia and Public Health Practice. Practice Guidelines for Public Health: Assessment of scientific Evidence, Feasibility, and Benefits: A report of the Guideline Development Project for Public Health Practice. Council on Linkages between Academia and Public Health Practice; 1995.
- Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses and randomized control trials and recommendations of clinical experts. *JAMA* 1992;268:240-8.
- Auston I, Cahn MA, Selden CR. Literature Search Methods for the Development of Clinical Practice Guidelines. In: McCormick KA, Moore SR, Siegel RA (eds.). *Clinical Practice Guideline Development: Methodology Perspectives*. U.S. Department of Health and Human Services, Agency for Health Care Policy and Research. Rockville, MD: AHCPR Pub. No. 95-0009, 1994, pp.123-127.
- Bryson JM. A strategic planning process for public and non-profit organizations. In: Bryson JM (ed.). *Strategic Planning for Public Service and Non-Profit Organizations*. Tarrytown, NY: Pergamon Press, 1993, pp. 11-19.
- Corber SJ. Community Health Practice Guidelines: An Important First Step. *Can J Public Health* 1994;85(Suppl 1): S5
- CDC. Counseling Practices of Primary-care Physicians - North Carolina, 1991. *MMWR* 1992;41:565-568.
- CDC. Assessing the public health threat associated with waterborne cryptosporidiosis. *MMWR* 1995;44(No. RR-6).
- Coughlin SS, Beauchamp TL (eds). *Ethics in Epidemiology*. Oxford University Press, 1996.
- Delbecq AL, Van de Ven AH and Gustafson DG. *Group Techniques for Program Planning: A Guide to Nominal Group and Delphi Processes*. Glenview, IL: Scott, Foresman, and Company, 1975.
- Eddy DM. Clinical decision making: from theory to practice. Guidelines for policy statements: the explicit approach. *JAMA*. 263(16):2239-40
- Emerson JD, Burdick E, Hoaglin DC, Mosteller F, Chalmers TC. An empirical study of the possible relation of treatment differences to quality scores in controlled randomized clinical trials. *Controlled Clin Trials* 1990;11:339-52.
- Eysenck HJ. Meta-analysis and its problems, *BMJ* 1994;309:789-92.
- Field MJ, Lohr KN (eds.). *Clinical Practice Guidelines: Directions for a New Program*. Institute of Medicine (U.S.). Committee to Advise the Public Health Service on Clinical Practice Guidelines. Washington, DC: National Academy Press, 1990.

- Field MJ, Lohr KN (eds.). *Guidelines for Clinical Practice: From Development to Use*. Institute of Medicine (U.S.). (Chapter 2: Developing Clinical Practice Guidelines) Committee on Clinical Practice Guidelines, Division of Health Care Services. Washington, DC: National Academy Press, 1992, pp 45-64. Chapter 8: A Framework for the Future, pp 196-219. Appendix A: Examples of Clinical Practice Guidelines and Related Materials, pp 243-345.
- Field MJ, Lohr KN (eds.). *Guidelines for Clinical Practice: From Development to Use*. Institute of Medicine (U.S.). [Chapter 8: A Framework for the Future. In:] Committee on Clinical Practice Guidelines, Division of Health Care Services. Washington, DC: National Academy Press, 1992, pp 196-219. Chapter 2: Developing Clinical Practice Guidelines, pp 45-64.
- Field MJ (ed.). *Setting Priorities for Clinical Practice Guidelines*. Institute of Medicine (U.S.). Committee on Methods for Setting Priorities for Guidelines Development, Division of Health Care Services. Washington, DC: National Academy Press, 1995.
- Green LW, Kreuter MW. Applications in Health-Care Settings: Changing the Behavior of Health Care Staff. In *Health Promotion Planning: An Educational and Environmental Approach*. Mayfield Publishing Company. Mountain View, California, 1991.
- Green LW, Kreuter MW. Social Diagnosis: Assessing Quality-of-Life Concerns. In *Health Promotion Planning: An Educational and Environmental Approach*. Mayfield Publishing Company. Mountain View, California, 1991.
- Greenland S. A critical look at some popular meta-analytic methods. *Am J Epidemiol* 1994;140:290-6.
- Greenland S. Quantitative methods in the review of epidemiologic literature. *Epidemiol Rev* 1987;9:1-30.
- Gyorkos TW, Tannenbaum TN, Abrahamowicz M, Oxman AD, Scott EA, Millson ME, et. al. An Approach to the Development of Practice Guidelines for Community Health Interventions. *Can J Public Health* 1994;85(Suppl 1): S8-S13.
- Gyorkos TW, Franco ED, Tannenbaum TN, Abrahamowicz M, Bedard L, Carsley J, et. al. Practice Survey of Immunization in Canada. *Can J Public Health* 1994;85(Suppl 1): S31-S36.
- Hahn R. Ethical Issues. In: *Principles and Practice of Public Health Surveillance*. Oxford University Press, Inc., New York, 1994.
- Hedges LV, Olkin I. *Statistical methods for meta-analysis*. Orlando, Florida: Academic Press, 1985.
- Henderson WG, Moritz T, Goldman S, Copeland J, Sethi G. Use of cumulative meta-analysis in the design, monitoring, and final analysis of a clinical trial: a case study. *Controlled Clinical Trials* 1995;16:331-41.
- Hugentobler MK, Israel BA, Schurman SJ. An Action Research Approach to Workplace Health: Integrating Methods. *Health Education Quarterly* 1992;19:55-76.

- Krumholz HM, Radford MJ, Ellerbeck EF, et al. Aspirin for secondary prevention of acute myocardial infarction in the elderly: prescribed use and outcomes. *Ann Intern Med* 1996;124:292-298.
- Lau J, Antman EM, Jimenez-Silva J, Kupelnick, Mosteller F, Chalmers TC. Cumulative Meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992; 327:248-54.
- Leape LL, Park RE, Solomon DH, et al, Group judgments of the appropriateness: The effect of a panel composition. *Quality Assurance in Health Care* 4:151-159, 1992.
- Linstone HA, Turoff M. (Eds.) *The Delphi Method: Techniques and Applications*. Adison-Wesley Publishing Co., Reading, MA. 1975.
- Lomas J, Anderson GM, Enkin M et al. "The Role of Evidence in the Consensus Process: results from a Canadian consensus exercise." *JAMA* 259(20):3001-3005, 1988.
- National Committee for Quality Assurance. HEDIS 3.0: Health Plan Employer Data and Information set. Developed Under the Auspices of the Committee On Performance Measurement, July 1996 (Draft for Public Comment).
- National Performance Review, Accompanying Report, September 1993.
- National Institutes of Health. Consensus Development Conference Program. Draft - Physical Activity and Cardiovascular Health. December 18-20, 1995. (Internet citation).
- Olkin I. Meta-analysis: reconciling the results of independent studies. *Stat in Medicine* 1995;14 (5/6/7):457-72.
- Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991;11:1271-8.
- O'Neill, RT. The biostatistical regulatory model. In *Clinical Practice Guidelines Development: Methodology Perspectives*. Eds. McCormick, KA, Moore, RS, Siegel, RA. U.S. DHHS, PHS, AHCPR. Pub. No. 95-0009; 1994.
- Field, MJ, Lohr, KN, Eds. *Guidelines for Clinical Practice: From development to Use*. Institute of Medicine (U.S). Committee on Clinical Practice Guidelines. National Academy Press, Washington, D.C. 1992. PHS Sponsored Guidelines Report Outline. July 26, 1995.
- Pettiti DB. *Meta-analysis, decision analysis, and cost-effective analysis: methods for quantitative synthesis in medicine*. New York: Oxford University Press, 1994.
- Pritzker DM and Dalton DS. *Negotiated Rule Making Sourcebook*. Office of the Chairman, Administrative Conference of the United States, 1995.
- Pritzker DM. The Blossoming of Reg-Neg. *National Institute for Dispute Resolution News*, Volume II, No. 4 July/August, 1995.
- Moore, CM. *Group Techniques for Idea Building*, Second Ed. Thousand Oaks: Sage Publications, 1994.

- Schriger DL. Training Panels in Methodology. In *Clinical Practice Guideline Development: Methodology Perspectives*. Eds. McCormick, KA, Moore, RS, Siegel, RA. U.S. DHHS, PHS, AHCPR. Pub. No. 95-0009; 1994.
- Snider DE, Holtgrave DR, Dunet DO. Decision Analysis. In: Haddix AC, Teutsch SM, Shaffer PA, Dunet DO. *Prevention Effectiveness: A guide to decision analysis and economic evaluation*. Oxford University Press, New York, 1996.
- Steckler A, McLeroy KR, Goodman RM, Bird ST, McCormick L. Toward Integrating Qualitative and Quantitative Methods: An Introduction. *Health Education Quarterly* 1992;19:1-8.
- Steinberg KK, Thacker SB, Smith SJ, et al. A meta-analysis of the effect of estrogen replacement therapy on the risk of breast cancer. *JAMA* 1991;265:1985-90.
- Sterling TD, Rosenbaum, WL, Weinkam JJ. Publication decisions revisited: the effect on the outcome of statistical tests on the decision to publish and vice versa. *Am Statistician* 1995;49:108-12.
- Thacker SB. Meta-analysis: a quantitative approach to research integration. *JAMA* 1988;259:1685-9.
- The Steering Committee of the Physician's Health Study Research Group. Preliminary report: findings from the aspirin component of the ongoing Physician's Health Study. *N Engl J Med* 1988;318:262-264.
- Thissen WHA, Sage AP, Warfield NJ. A user's guide to public systems methodology. Charlottesville, VA: School of Engineering and Applied Science, 1980.
- U.S. Preventive Services Task Force. Guide to Clinical Preventive Services: An Assessment of 169 Interventions. Report of the U.S. Preventive Service Task Force. Baltimore, MD: Williams and Wilkins, 1989, pp.375-379.
- U.S. Preventive Services Task Force. Guide to clinical preventive services, 2nd ed. Baltimore: Williams & Wilkins, 1996.
- U.S. Congress, Office of Technology Assessment, Identifying Health Technologies that Work: Searching for Evidence, OTA-H-608 (Washington, DC: U.S. Government Printing Office, September, 1994). Chapter 8: The Impact of Clinical Guidelines on Practice. In Warfield NJ. *Societal systems: Planning, policy, and complexity*. New York: John Wiley, 1976.
- Weed DL. Science, Ethics Guidelines, and Advocacy in Epidemiology. *Ann Epidemiol* 1994;4:166-171.
- Weingarten S, Stone E, Hayward R, et al. The adoption of preventive care practices by primary care physicians. Do actions match intentions? *J Gen Intern Med* 1995; 10:138-144.
- Wolcott HF. Posturing in Qualitative Research. In: *The Handbook of Qualitative Research in Education*. Academic Press, Inc., Harcourt Brace & Company, San Diego, 1992.

- Woolf SH. An organized analytic framework for practice guideline development: Using the analytic logic as a guide for reviewing evidence, developing recommendations, and explaining the rationale. In: McCormick KA, Moore SR, Siegel RA (eds.). *Clinical Practice Guideline Development: Methodology Perspectives*. U.S. Department of Health and Human Services, Agency for Health Care Policy and Research. Rockville, MD: AHCPR Pub. No. 95-0009, 1994, pp.105-113.
- Woolf SH. Practice Guidelines, a New Reality in Medicine: III. Impact on Patient Care. *Arc Intern Med* 1993;153:2646-2655.
- Woolf SH. Practice Guidelines: a New Reality in Medicine II: Methods of Developing Guidelines. *Arc Intern Med* 1992;152:946-952.

This page intentionally left blank.

Glossary

Administrative issues. Issues of resource mobilization and process control complementary to the central scientific tasks of guideline development.

Advocacy. Working for political, regulatory, or organizational change on behalf of a particular interest group or population.

Alternative approaches. Examples of different ways of accomplishing the same guideline development task.

Analytic framework. A flow chart, influence diagram, decision tree, or clinical algorithm used to describe or illustrate the chain of causal reasoning which links a recommended health practice to the desirable health outcomes in a defined individual or population by means of credible evidence of effectiveness.

Archival research. Examining reports, newspaper clippings, correspondence, books, and other documents prepared by persons other than the researcher and kept in archives (e.g., libraries). Archival sources can reveal changes in practitioner perceptions and practices over time.

Assessment. Estimation of the relative magnitude, importance, or value of objects observed.

Attitude. A relatively constant feeling, predisposition, or set of beliefs directed toward an object, person, or situation.

Behavior. An action that has a specific frequency, duration, and purpose, whether conscious or unconscious.

Belief. A statement or proposition, declared or implied, that is emotionally and/or intellectually accepted as true by a person or group.

Benefits. Valued health outcomes or improvements in quality of life or social conditions having some known relationship to health promotion or health-care interventions.

Clinical preventive service. A primary care clinical practice or preventive intervention—screening test, counseling intervention, immunization, or chemoprophylactic regimen—for target conditions among asymptomatic individuals of all age groups and risk categories.

Community. A collective of people identified by common values and mutual concern for the development and well-being of their group or geographical area.

Community forums or workshops. Intensive meetings of interest groups assembled to discuss issues surrounding proposed guidelines and to reach an understanding among the participants about needs, scope, format, timetables, and roles.

Core functions of public health agencies. The unique functions of public health agencies at all levels of government: assessment, policy development, and assurance.

Assessment. Regular and systematic collection, assembly, analysis, and dissemination of information on the health of the community, including statistics on health status, health needs, and epidemiologic and other studies of health problems.

Policy development. Promoting the use of the scientific knowledge base in decision-making about public health, leading in developing public health policy, and taking a strategic planning approach to influencing the democratic political process to serve the public interest in good health.

Assurance. Assuring agency constituencies that services necessary to achieve agreed upon goals are provided, either by encouraging actions by other entities (private or public), by requiring such action through regulation, or by providing services directly.

Cost-effectiveness. A measure of the cost of an intervention relative to its population impact, usually expressed in dollars per unit of health outcome achieved.

Decision analysis. An explicit, quantitative, systematic approach to decision making under conditions of uncertainty.

Delphi technique. A method of sampling the opinions or preferences of a small number of experts, opinion leaders, or informants, whereby successive questionnaires are sent by mail and the results (rankings or value estimates) are summarized for further refinement on subsequent mailings.

Desirable attributes of guidelines. Quality characteristics which help to persuade users that guidelines are trustworthy, easy to put into practice, and effective in achieving desired health outcomes.

validity Supported by strong evidence linking recommendations to outcomes.

reliability or reproducibility Prepared using procedures and decision rules that would lead different experts to the same conclusions based on the same evidence.

applicability [in practice] Useful in populations that potential users would consider relevant to their practices.

flexibility [in practice] Allowing for practitioner judgment and patient preferences.

clarity Presented in unambiguous language and easy-to-follow logic.

multidisciplinary process Prepared with input from relevant disciplines and stakeholders.

up to date [scheduled review] Reflecting the most recent evidence and including a written timetable for future revision.

well documented Published along with explicit statements on assumptions, process, rationale, evidence, and decision rules.

Effectiveness. The improvement in health outcome that a prevention strategy can produce in typical community-based settings.

Efficacy. The improvement in health outcome that a prevention strategy can produce in expert hands under ideal circumstances.

Emerging health problem. Recently discovered or re-discovered public health problem such as such as hantavirus pulmonary syndrome or multidrug-resistant tuberculosis.

Enabling factors. Necessary skills and resources (finances, staff, space, and educational materials) which the professional must possess in order to successfully implement the guideline's recommendations.

Essential public health services A list of ten categories of public health activities which the US Public Health Service considers essential for maintaining optimal community health status.

Monitoring **health status** to identify community health problems.

Diagnosing and investigating **health problems and hazards** in the community.

Informing, educating, and empowering people about **health issues**.

Mobilizing community **partnerships** to identify and solve health problems.

Developing **policies and plans** that support individual and community health efforts.

Enforcing **laws and regulations** that protect health and ensure safety.

Linking people to needed **personal health services** and assuring the provision of health care when otherwise available.

Assuring a competent public health and personal health care **workforce**.

Evaluating effectiveness, accessibility, and quality of personal and population-based health services.

Researching for **new insights** and innovative solutions to health problems.

Focus group interviews. Informal group interviews in which eight to 12 potential users of a proposed guideline are asked to discuss their thoughts and feelings about the issue. Those thoughts and feelings are then used to help clarify the content, delivery, and appeal of the proposed guidelines.

Group process methods. Formal techniques for maximizing the contributions of all participants in the guideline development process.

Guideline development method. A set of rules of procedure for collecting evidence of effectiveness of a health practice and for making group decisions about the quality and sufficiency of that evidence in supporting specific practice recommendations.

Informal consensus. Participants on an expert panel decide on what to recommend on the basis of open discussion during one or more meetings. Because the process is informal, the recommendations are often not accompanied by explanations of the causal reasoning which links the recommended professional behavior, credible evidence of effectiveness, and the desirable health outcomes.

Formal consensus. Participants on an expert panel decide on what to recommend by consensus during the course of a structured two and one-half day conference. The meetings

provide greater structure to the analytic process than the informal consensus method. Consensus statements do not usually include references to the literature, nor the underlying rationale or evidence behind any recommendations or conclusions.

Evidence-based. The evidence-based method aims to explicitly link practice recommendations to the underlying scientific evidence of effectiveness.

Explicit method. A subcategory of the evidence-based method which relies on mathematical modeling and other formal analytic methods to generate estimates of the probability of occurrence of specific benefits, harms, and costs of alternative practices.

Guideline development task. An activity or piece of work that results in a product needed to plan, develop, disseminate, and encourage users to adopt a guideline. The CDC guideline development process is divided into 13 discreet tasks.

Planning and coordinating the process. Defining objectives, mobilizing resources, and overseeing activities to ensure timely and efficient achievement of defined objectives.

Assessing user needs. Finding out about the nature, extent, and determinants of current practices regarding the candidate intervention, technology, or health problem of interest in the population of primary guideline users; and gauging the level of common concern and consensus on the potential utility of the proposed guideline.

Choosing guideline topics. Deciding on the topics that should be given priority for guideline development and explaining the rationale for each choice.

Selecting guideline panels. Choosing the members and leaders of the panel of experts and other participants who will assess the scientific evidence and formulate guideline recommendations.

Defining the scope of guidelines. Delimiting the target population, outcomes, and interventions which are (1) of greatest interest to specific practitioners, the public, and other users, (2) most amenable to clarification by means of systematic assessment and synthesis of scientific evidence, and (3) capable of being fully explored and resolved into clear and specific advice within the limitations of time and resources.

Clarifying the method and analytic framework. Justifying the use of one or more guideline devel-

opment methods (informal consensus, formal consensus, evidence-based, or explicit) and describing or illustrating the chain of causal reasoning which links the recommended health practice to the desirable health outcomes in a defined individual or population by means of credible evidence of effectiveness.

Identifying and synthesizing the evidence. Seeking, collecting, and assessing the quality and quantity of empirical evidence of the effectiveness of a proposed intervention for ensuring a desirable health outcome in a defined population. Empirical evidence of effectiveness must be systematically identified, synthesized, and documented using methods that minimize bias and maximize precision.

Aiding group interaction and decision making. Using formal techniques to maximize the contributions of all participants in the guideline development process. Group interaction methods add clarity and explicitness to group decision making, balance power among participants, minimize bias, and ensures documentation of decision rules and products.

Identifying a research agenda. Identifying a list of study questions or areas of inquiry that should receive high priority for scientific investigation and funding. The primary source of study questions is the analytic framework, its assumptions, and presumed causal linkages for which existing evidence is inadequate.

Updating the guideline. Deciding on a timetable for revising the guideline to reflect new scientific knowledge. The timetable is determined by the strength of current supporting evidence and expectations about new discoveries, effectiveness of the guideline, and changes in the practice environment.

Writing the guideline. Preparing a document written with unambiguous language and easy-to-follow logic that provide clear recommendations and document the rationale on which the recommendations are based.

Obtaining critical reviews and public comment. Obtaining input from content and policy experts, practitioners, advocates, and the public about the scientific accuracy, completeness, and ease of implementation of the draft guidelines. Revising the draft guideline in response to such feedback from potential users and other interested persons can increase the credibility of the guidelines.

Encouraging adoption of the guideline. Ensuring that the guideline is appealing, widely disseminated, and encourages potential users to accept and implement the recommended practices.

Health outcome. A medically or epidemiologically defined characteristic of a patient or health problem in a population that results from health promotion or care provided or required as measured at one point in time.

Health promotion. A planned combination of educational, political, regulatory, and organizational supports for actions and conditions of living conducive to the health of individuals, groups, or communities.

Idea writing. This technique promotes group interaction in large groups by writing rather than by discussion. The large group is divided into working groups of 4 to 5 individuals; participants react in writing to a stimulus question or item and then place their pad (with the initial response) in the center of the group; each participant reacts, in writing, to what is written on each of the other pads; participants read the comments made in reaction to their initial response; the small working groups discuss the principal ideas that emerge from the written interaction; and the group summarizes the discussion in writing.

Implementation. The act of converting guideline recommendations into actions through policy, regulation, and organizational incentives.

Influence diagrams. Diagrams of factors that are important in a decision and the interrelationships between factors in a model resembling a flow chart. Mathematical calculations can be performed for models if numerical values are assigned to designate the probability of certain events happening and the utility, or perceived usefulness, of each possible outcome.

Intervention. The part of a health strategy, incorporating method and technique, that actually reaches a person or population.

Key informant interviews. In-depth, semi-structured interviews with individuals who are selected to participate because they have special knowledge or insight not available to others in their reference group—hence the label “key informant.” Recorded or written transcripts of the interviews are then analyzed to uncover themes and the rela-

tive importance of frequently mentioned issues relating to guideline development and use.

Meta-analysis. Data from multiple studies are synthesized to give a statistical measure (e.g., an odds ratio or relative risk) of effect and an assessment of its significance. The combination of data from several studies increases both the statistical power and ability to generalize.

Narrative literature reviews. Expert opinion papers, in journals and text books, which lack systematic rules for acquiring or consolidating the evidence. In a traditional narrative review of the medical literature, a subject-matter expert reviews studies, decides which are relevant to a particular topic, and highlights findings in terms of results and, to a lesser degree, methodology.

Need (professional). An estimate of the knowledge, skills, motivation, and incentives that a practitioner requires to encourage use of a proposed guideline.

Negotiated rule making techniques. A consensus-based approach in which a committee of affected interested parties (agency regulators and the regulated parties), meets with a skilled mediator. The goal is to reach consensus on the proposed policy. Negotiations that do not result in consensus can also be useful by narrowing issues in dispute, identifying information necessary to resolve disputes, ranking priorities, finding potentially acceptable solutions, and improving the agency’s understanding of the real-world impact of alternative options.

Nominal group process or technique. A technique which elicits written responses to a single question without verbal interaction among group members. The process would include small groups of five to nine potential guideline users to assess target group perceptions of need and obstacles to meeting that need. It reduces the tendency for the more socially powerful to dominate the discussions and to bias the consensus that emerges.

Planning. The process of defining needs, establishing priorities, diagnosing causes of problems, assessing resources and barriers, and allocating resources to achieve objectives.

Policy. Objectives and rules guiding the activities of an organization and providing authority for allocation of resources.

Predisposing factors. Professional's values, beliefs, attitudes, and perceptions about guidelines and their potential usefulness. These factors account for the professional's motivation to use guidelines and confidence in being able to implement their recommendations.

Prevention effectiveness. A systematic effort to assess the impact of public health policies, programs, and practices on costs and health health outcomes.

Prevention program. A set of planned activities over time designed to achieve specified objectives in terms of health problems prevented.

Program evaluation. An assessment of the processes, impacts, and outcomes of program activities in relation to the objectives, standards of acceptability, and expectations of stakeholders.

Public health practice guideline. A systematically developed statement which helps policy-makers, public health practitioners, clinical practitioners, health agencies, and the public decide on appropriate actions to (1) promote health, (2) prevent disease, disability, and premature death, and (3) improve quality of life of members of a defined population. Appropriate actions may target individuals or whole communities at risk for adverse health outcomes.

Public health practice. Organized community effort (public and private) to address the public interest in health by applying scientific and technical knowledge to prevent disease and promote health.

Qualitative methods for assessing needs. Techniques adopted from the social sciences to elicit an "insider's view" (study participants or participant observers) of how members of a group under study perceive their own needs for guidelines, how they are likely to react to such guidelines, and what positive or negative consequences they expect to result from the use of those guidelines.

Quality assessment. Measurement of professional or technical practice or service for comparison with accepted standards to determine the degree of excellence.

Quality assurance. Formal process of implementing quality assessment and quality improvement in

programs to assure stakeholders that professional activities have been performed appropriately.

Quality of life. The perception of individuals or groups that their needs are being satisfied and that they are not being denied opportunities to achieve happiness and fulfillment.

Quantitative methods for assessing needs. Numerical measurements and statistical techniques that estimate aggregate characteristics of the group under study and support inferences which can be generalized to a larger reference population. Such methods include sample surveys and synthetic estimates from administrative data sources.

Reinforcing factors. Rewards or incentives that are anticipated or that actually follow as a consequence of a particular behavior. They include reimbursements actually received, visible improvements in patient or population health outcomes, support from colleagues, and feedback from patients or clients.

Research agenda. A list of study questions or areas of inquiry that should receive high priority for scientific investigation and funding.

Self-efficacy. A construct from social learning theory referring to the belief an individual holds that he or she is capable of performing a specific behavior.

Stakeholders. People who have an investment or stake in the outcome of an intervention or program and therefore have reasons to be interested in the evaluation of the program.

Strategy. A plan of action that anticipates barriers and resources in relation to achieving a specific objective.

Systematic literature reviews. Reviews based on rules to identify studies and collate information from studies. Such reviews are often the initial step in meta-analyses.

Value. A preference shared and transmitted within a community.

This page intentionally left blank.

Abbreviations

ACET	Advisory Committee for the Elimination of Tuberculosis
ACIP	Advisory Committee for Immunization Practices
ADS	Associate Director for Science
AHCPR	Agency for Health Care Policy Research
ALA	American Lung Association
ATS	American Thoracic Society
ATSDR	Agency for Toxic Substances Disease Registry
CBA	Cost Benefit Analysis
CBO	Community Based Organization
CDC	Centers for Disease Control and Prevention
CDCP	Consensus Development Conference Program (NIH)
CEA	Cost Effectiveness Analysis
CFR	Code of Federal Regulations
CHPG	Canadian Community Health Practice Guidelines
CIO	Center, Institute, or Office
CQI	Continuous Quality Improvement
DTBE	Division of Tuberculosis Elimination
EISC	Excellence in Science Committee
EPO	Epidemiology Program Office
FACA	Federal Advisory Committee Act
FDA	Food and Drug Administration
FR	Federal Register
GDPPHP	Guideline Development Project for Public Health Practice
GMWG	Guidelines Methodology Working Group
HEDIS	Health Plan Employer Data and Information Set
HICPAC	Hospital Infections Control Practices Advisory Committee
HIP	Hospital Infections Program
HIV	Human Immunodeficiency Virus
HMO	Health Maintenance Organization
IDSA	Infectious Diseases Society of America
IOM	Institute of Medicine
MCC	Office of the CDC Managed Care Coordinator
MCO	Managed Care Organization
MMWR	Morbidity and Mortality Weekly Report
NCCDPHP	National Center for Chronic Disease Prevention and Health Promotion
NCDEHNR	North Carolina Department of Health, Environment, and Natural Resources
NCEH	National Center for Environmental Health
NCHSTP	National Center for HIV, STD, and TB Prevention
NCI	National Cancer Institute
NCID	National Center for Infectious Diseases
NCIPC	National Center for Injury Prevention and Control
NGT	Nominal Group Technique
NIH	National Institutes of Health
NIOSH	National Institute for Occupational Safety and Health
NIP	National Immunization Program
OD	Office of the Director
OGC	Office of General Counsel
OMAR	Office of Medical Applications of Research (NIH)
OPPE	Office of Program Planning and Evaluation
OTA	Office of Technology Assessment (US Congress)
PEA	Prevention Effectiveness Activity
PHPPPO	Public Health Practice Program Office

PHS	Public Health Service
PSA	Prostate specific antigen
RCT	Randomized Controlled Trials
RR	Reports and Recommendations (MMWR)
STD	Sexually Transmitted Diseases
TB	Tuberculosis
USDHHS	United States Department of Health and Human Services
USPHS	United States Public Health Service
USPSTF	United States Preventive Services Task Force
WWW	World Wide Web

APPENDIX A: DESIRABLE ATTRIBUTES OF CLINICAL PRACTICE GUIDELINES¹

Attribute	Explanation
VALIDITY	Practice guidelines are valid if, when followed, they lead to the health outcomes projected for them. A prospective assessment of validity will consider the substance and quality of the evidence cited, the means used to evaluate the evidence, and the relationship between the evidence and recommendations.
Strength of Evidence	Practice guidelines should be accompanied by descriptions of the strength of the evidence and expert judgment behind them.
Estimated Outcomes	Practice guidelines should be accompanied by estimates of the health and cost outcomes expected from the interventions in question, compared with alternative practices. Assessments of relevant health outcomes will consider patient perceptions and preferences.
RELIABILITY/ REPRODUCIBILITY	Practice guidelines are reproducible and reliable (1) if---given the same evidence and methods for guidelines development---another set of experts produces essentially the same statements and (2) If---given the same clinical circumstances---the guidelines are interpreted and applied consistently by practitioners (or other appropriate parties).
CLINICAL APPLICABILITY	Practice guidelines should be inclusive of appropriately defined patient populations as evidence and expert judgment permit, and they should explicitly state the population(s) to which statements apply.

¹ IOM (Institute of Medicine). Summary. Table 1. In *Guidelines for Clinical Practice: From Development to Use*. M.J. Field and K.N. Lohr, editors. Washington, DC: National Academy Press, 1992.

Attribute	Explanation
CLINICAL FLEXIBILITY	Practice guidelines should identify the specifically known or generally expected exceptions to their recommendations and discuss how patient preferences are to be identified and considered.
CLARITY	Practice guidelines must use unambiguous language, define terms precisely, and use logical and easy-to-follow modes of presentation.
MULTIDISCIPLINARY PROCESS	Practice guidelines must be developed by a process that includes participation by representatives of key affected groups. Participation may include serving on panels that develop guidelines, providing evidence and viewpoints to the panels, and reviewing draft guidelines.
SCHEDULED REVIEW	Practice guidelines must include statements about when they should be reviewed to determine whether revisions are warranted, given new clinical evidence or professional consensus (or lack of it).
DOCUMENTATION	The procedures followed in developing guidelines, the participants involved, the evidence used, the assumptions and rationales accepted, and the analytic methods employed must be meticulously documented and described.

APPENDIX B: PROVISIONAL INSTRUMENT FOR ASSESSING CLINICAL PRACTICE GUIDELINES¹

KATHLEEN N. LOHR AND MARILYN J. FIELD

Division of Health Care Services

The Institute of Medicine (IOM) of the National Academy of Sciences has been engaged since the beginning of 1990 in two projects relating to the development, implementation, and evaluation of clinical practice guidelines. One IOM committee defined practice guidelines as "systematically developed statements to assist practitioners and patients in choosing appropriate health care for specific clinical conditions." It also delineated several desirable attributes of guidelines that are intended to help users understand the elements of a sound guideline and to recognize good (or not-so-good) guidelines.

These aspects of guidelines were discussed in *Clinical Practice Guidelines: Directions for a New Program* (IOM, 1990) and *Guidelines for Clinical Practice: From Development to Use* (this report).

The first IOM study committee discovered that no explicit method was available for assessing existing or emerging practice guidelines. At least one instrument was being tested to assess some aspects of guideline development (AMA, 1990), but nothing existed to judge the quality, reliability and validity of the content of the guideline itself. Therefore, one task the second IOM committee undertook was to develop an "assessment instrument" that could be used by various parties in formal evaluations of guidelines.

The next sections of this document describe, first, the purposes of the "provisional" assessment instrument and, second, its development. The discussion covers several features of the instrument and its application, notes several cautions and caveats about the present form, all of which warrant further consideration. Finally, the document presents the instrument itself, in three operational parts—a general information sheet, the full instrument, and a summary evaluation sheet. The instrument is termed provisional because the committee firmly believed that more experience needs to be accumulated by testing it on different kinds of guidelines.

PURPOSES OF THE ASSESSMENT INSTRUMENT

The central purpose of the IOM's instrument for assessing clinical practice guidelines is to provide an explicit method for examining the soundness of such guidelines and to encourage their systematic development. By assessment is meant a prospective judgment of the soundness of both the process used in developing a guideline and the resulting guideline. The intent is to avoid situations in which a guideline that is not

¹IOM (Institute of Medicine). Appendix B. In *Guidelines for Clinical Practice: From Development to Use*. M.J. Field and K.N. Lohr, editors. Washington, DC: National Academy Press, 1992.

consistent with the scientific evidence is nonetheless "rated" as good on procedural criteria alone.²

More concretely, the IOM intended to operationalize its attributes of good guidelines and to provide a standardized approach and structure for the assessment of a guideline document. The resulting form is not simple. Therefore, the IOM does not expect practicing physicians or other clinicians, patients, other nonprofessionals, or policymakers to apply this instrument. Rather it expects individuals (or groups) with three types of expertise to apply it—namely, those with clinical experience with patients who have the conditions or problems covered by the guideline document, those with research experience in the conditions or technologies covered, and those with methodologic skills in developing guidelines. Any final or overall judgments of a guideline document emerging from the application of this instrument would be reported in simpler, summary form in ways that would convey the relative soundness (or lack of it) of a given guideline document to all potential users of the guideline.

The IOM committee sees three possible uses of this instrument: as an educational tool, as a self-assessment tool, and as a means of judging guidelines before their adoption. The Agency for Health Care Policy and Research (AHCPR) may want to use this instrument, or one like it, in directing the work of its guidelines-development panels; the agency may also wish to employ it in judging the products of those panels or the guidelines developed by other groups, such as medical specialty societies. Furthermore, other groups may wish to review existing or draft guidelines of their own against this instrument, in an effort to identify guidelines warranting revision or defects in draft guidelines that warrant correction before they are put into final form. Finally, if an organization were to be created for the express purpose of certifying or otherwise reporting on the soundness of particular guidelines, it might wish to employ the instrument as part of its review activities.

DEVELOPMENT PROCESS

The concept of the instrument originated during discussions with members of the AHCPR staff about their responsibilities for practice guidelines under the Omnibus Budget Reconciliation Act of 1989. In seeking to respond to those early ideas, the IOM committee went through four steps.

First, with the aid of an outside consultant, staff drafted a set of questions to operationalize the eight conceptual attributes of good practice guidelines identified in the IOM's 1990 report on guidelines.³ (Discussions of these attributes introduce sections of the assessment instrument.) Second, these questions were combined with background information and instructions for users and subjected to considerable internal and external

²Evaluation of the eventual impact of guidelines is a separate step. Both IOM reports (1990 and this one) include discussions of evaluation.

³Anne-Marie Audet, M.D., of the Health Institute. New England Medical Center. Inc., served as consultant during the initial stage of the project.

review, as described below. Members of the IOM committee twice reviewed drafts of the assessment instrument during this time.

Third, IOM staff used the critiques and suggestions of reviewers to revise the instrument, and it, together with background material, was subjected to external review according to IOM and National Research Council (NRC) procedures. Fourth, the instrument was revised in response to that external review, resulting in the provisional questionnaire and other forms incorporated into this document.

Initial Reviews

An interim draft of the instrument was sent to AHCPR in December 1990 and was subsequently forwarded to several professional societies that had volunteered to review and test the instrument against guidelines of their own. By June 1991, the committee had received more than 15 separate responses and commentaries. Responses included a lengthy summary review provided by IMCARE (Internal Medicine Center to Advance Research and Education), which had solicited reviews from the 231 internists then in its Guideline Network. IMCARE also sent the document to 147 physicians who requested more information and received 65 responses from network members who reviewed and, in many cases, applied the instruments to real guidelines.

Reactions to the draft instrument were extremely varied. With respect to format, instructions, ease of use, and similar issues, positive comments included the following:

- "Well written-concise."
- "The instrument was easy to apply and the definitions of attributes were helpful in driving the assessment."
- "The meaning of each attribute, the instructions, and response categories were generally clear."
- "Overall, the instrument provides a precise algorithm for examining the degree to which a clinical guideline meets defined aspects of seven attributes. In our opinion, the instrument can be used for its intended purpose. In addition, application of an instrument such as this is, in itself, a thought-provoking exercise for individuals active in developing clinical practice guidelines."

By contrast, negative comments were of the following kind:

- "The form seems excessively lengthy and not particularly user-friendly."
- "The instructions are verbose and redundant, and almost legalistic. They are not user-friendly."
- "This instrument was very confusing, almost 'impossible.'"
- "Instrument of intellectual torture. Beyond Bureaucracy! . . . This makes me feel stupid!"

A form that accompanied the instrument asked reviewers to indicate the time (in person-hours) needed to apply the assessment instrument to a guideline of their choice. Among 59 individuals in the IMCARE group who evidently applied the instrument to an actual guideline and completed the form, 13 said that the instrument took under two hours to apply, 38 said two to three hours, and 8 said four hours or more. Several commentators indicated that their learning curve was quite steep and that repeated use of the instrument would make it simpler and less time-consuming to apply.

In rating the overall difficulty of understanding or using the assessment instrument, the majority of respondents indicated that it was moderately to very difficult. In addition, the great majority found the instrument good or at least somewhat helpful in helping them reach an overall judgment of the strengths and weaknesses of the guideline they were evaluating. Finally, more than half indicated that the instrument definitely should be revised; most of the rest were uncertain or had no opinion, and only about 1 in 10 advised abandoning the effort.

Among the more concrete recommendations for revisions were the following: (1) add more "don't know" or "not applicable" responses to certain questions, (2) simplify and shorten the instructions; (3) include a "prologue" containing pertinent information from the first IOM (1990) report on guidelines, which defines key terms and similar concepts; (4) consider adding an attribute related to endorsement by appropriate affiliated or outside organizations; and (5) clarify exactly who the intended users are. In addition many reviewers offered comments on the draft instrument itself, chiefly observations about the wording of questions and the addition of response categories. More general suggestions included employing the instrument to help generate guidelines (rather than rate them) and using the instrument to guide an assessment process (without necessarily requiring that the instrument be completed in full).

Final Reviews

In accordance with IOM and NRC procedures, the revised instrument was subjected to an external, anonymous review by a panel similar to the full IOM committee. The "provisional" form in this document reflects the reactions of this group of seven experts, and many of their comments have been incorporated into the discussion presented here.

IMPORTANT FEATURES OF THE PROVISIONAL ASSESSMENT INSTRUMENT

Attributes of Practice Guidelines

Types of Attributes

Four attributes identified in the first IOM report on practice guidelines concern the substance of the guidelines---clinical applicability or scope, clinical flexibility, reliability/reproducibility, and validity. Four others---clarity, multidisciplinary process,

scheduled review, and documentation---have more to do with process. This instrument explicitly incorporates all attributes but documentation; that attribute is captured in questions directed at the other seven. Each attribute is described in the text of the instrument.

Implicit Weight Accorded to Different Attributes

As discussed below, this instrument has no explicit or quantitative scoring system. The attributes are implicitly weighted, however, according to the number of main questions used to cover them. Of a total of 46 questions, validity has 22 questions; clarity, 8; multidisciplinary process, 4; clinical flexibility, 4; reliability and reproducibility, 4; clinical adaptability, 3; and scheduled review, 1. By this rough metric, validity is accorded by far the major emphasis in the document, reflecting the committee's concern for finding a way to judge the soundness of the guidelines themselves rather than just the acceptability of the development process.

Questions and Response Categories

Questions

Most of the seven attributes are dealt with through one or more questions that tap specific "dimensions." For example, validity is divided into five issues: (1) strength of the scientific evidence and professional consensus; (2) qualitative and quantitative statements about health benefits and harms or risks; (3) qualitative and quantitative statements about expected health costs or expenditures; (4) the extent to which specific recommendations are justified by the estimates of benefits, harms, and costs provided in the guidelines and the extent to which those estimates are supported by the evidence amassed in the guidelines document; and (5) potential conflicts among existing guidelines, if any.

The instrument has 46 descriptive questions related to the seven attributes. Generally they pertain to the presence of information about the attribute or about a particular dimension of an attribute. Several questions have additional "items" designed to help assessors think about key points implied by the main question and whether a particular dimension of an attribute is satisfactory or not.

Responses to Questions

Responses to each question are typically "yes" or "no" for questions asking about the presence or absence of certain information, features, or development processes. If the answer is "yes," a follow-up question asks about the quality of that information, feature, or development process---essentially whether the information provided is satisfactory or not. If the answer to the main question is "no," the follow-up question probes the significance

of the absence of information, a particular feature, or development process and asks whether the omission is important or not.

Satisfactory. Assessors can judge information about a particular attribute as satisfactory if all critical elements have been considered and presented. For example, the discussion or description should be thorough and comprehensive; the guideline developers should have based their work on appropriate and correct information (e.g., from the literature review); and they should have used appropriate methods (e.g., for evaluating the strength of the scientific evidence or reaching professional peer consensus).

Conditionally satisfactory. The description or discussion of an attribute or dimension is conditionally satisfactory if some, but not all, of the critical elements have been considered and presented. For example, the discussion of a particular aspect of an attribute such as scheduled review may be vague or incomplete. Alternatively, guidelines developers may have disregarded important information (about, for instance, likely risks to patients from the use of a technology) in reaching their recommendations, or they may have improperly used certain kinds of methods. These problems with the guideline document may not prevent a clinician from using it or understanding its recommendations, but they may affect its overall usefulness or call certain recommendations into question. Revisions would be presumed to improve the guidelines, but they would not be mandatory.

Unsatisfactory. The description or discussion about a specific attribute is unsatisfactory if most of the critical elements have not been considered or presented. For example, well-known pieces of clinical information (or the views of multiple specialists with an interest in the guideline topic) may have been ignored; methods of analysis may have been misapplied; or recommendations may be based on faulty information or poor logic, or both. In such a case, it would be difficult if not impossible to judge the quality of the process of guideline development or the soundness of the resulting guidelines and recommendations (or both); certainly assessors could not mark those attributes as satisfactory. Serious thought must be given to augmenting or revising the guidelines document before it is promoted further.

Omissions. Omitting a description or discussion about a specific attribute or dimension may be *unimportant* if that omission seems likely to have no demonstrable effect either on the ability of a guideline user to apply the guideline effectively in the clinical decision-making process or on the capacity of the assessor to make an independent assessment of the quality of that attribute. Omitting such a description or discussion is of *minor importance* if it seems likely to affect negatively the ability of a guideline user or of evaluators to apply the guideline effectively or independently to assess its quality. Finally, omitting such a description or discussion is a *major omission* if the absence of such information essentially prevents guideline users or evaluators from applying the guideline effectively or even making an independent evaluation about the soundness of the guidelines document itself (at least on that particular feature).

Special Cases

Special cases may arise in which information appropriately is omitted from the guideline because the question or item is not applicable or is inappropriate (given responses to earlier items, for instance). In such an instance, the assessor is asked to mark the response category most appropriate for the given case (e.g., not applicable). In other situations, assessors may find it difficult to arrive at a single answer to the question, especially if the guidelines document being evaluated is very complex or if necessary background information appears to be missing. "Comments" sections are provided throughout for assessors to record additional remarks or qualifying statements, to highlight areas not well covered by the instrument, and to note special factors that should either be followed up or taken into account in the overall judgment about the guidelines document. Finally, if assessors conclude that the guidelines document is so complex, clinically esoteric, or methodologically sophisticated that it warrants additional, outside expert review, they are asked to note that at the end of the full instrument and also on the summary evaluation sheet.

Alternative Approaches to Responses

The main type of response used in this form is categorical (e.g., satisfactory, conditionally satisfactory, and unsatisfactory). Some reviewers noted that this approach is inherently constraining and requires definitions of the three categories, which may not be interpreted consistently. Furthermore, these categories do not allow assessors to distinguish guidelines that more properly should be characterized as excellent or outstanding. To overcome some of these drawbacks, an approach to responses based on a scale might be tested.

For example, a five- or even seven-point scale might be adopted, with one end of the range described as excellent (exemplary, highly satisfactory, or a similar superlative) and the opposite end described as poor (inferior, or very unsatisfactory). The equivalent approach might also be tried for the responses concerned with omissions of information. The committee believes this change warrants testing at some point in the future development of this form.

Response Scoring

After considerable debate and consideration of reviewers' comments, the committee concluded that this instrument should not be "scored" in any quantitative way. Thus, it does not propose any formal weighting or numerical scoring scheme for the main questions, nor does it suggest a particular threshold, cutpoint, or floor against which current guidelines might be judged acceptable or unacceptable. If most responses to the questions are "satisfactory" (or "unimportant omissions"), however, one might reasonably conclude that such a guidelines document would be sufficient for most clinical situations. Alternatively, if most responses were unsatisfactory (or major omissions), one would probably argue that the guidelines document needed to be revised before it could be used effectively.

The committee was of the view that a defensible scoring system could not be designed *a priori* in any case, regardless of whether scoring would be purely categorical or more quantitative. Review and testing of the assessment instrument itself---with revisions as necessary---will be required before a sensible scoring system can be proposed. Moreover, different users of the assessment instrument may have legitimate reasons to differ on where they would establish such cutpoints. Provision of information on the quality of guidelines documents appears to be more in the public interest than is making "one-size-fits-all" judgments on behalf of others.

In the same vein, no single question is treated as the signal of a "fatal flaw." That is, for no question will a response of "no," "unsatisfactory," "major omission" *by itself* render the guidelines document unacceptable. Some committee members believed that certain questions, especially those relating to validity, should be so designated. However, the questions that seem to be the most likely candidates for this level of decisiveness⁴ were added in response to the external review of the draft document; therefore they have not yet been reviewed or tested further. The committee believes designating these (or other) items as potentially "fatal" is premature.

Response Aggregation and Display

General Comments

Information obtained from applying the instrument might eventually be arrayed in one or more qualitative, summary displays or tables, as might be done, for instance, by a *Consumer Report* article. This might provide a rough indicator of whether the guidelines document could be used effectively in clinical situations. One reviewer, for example, suggested that a report for busy practicing physicians might usefully include "a graphic summary of the degree to which each attribute was successfully achieved, e.g., a bar representing the percent of key items within each attribute that were deemed satisfactory . . . ; and . . . a brief, narrative summary assessment."

Summary Evaluation Sheet

The committee did not pursue the design of such displays, chiefly because such an effort was seen as premature for an instrument that itself warrants additional testing and application. As an intermediate step, however, the instrument does include a "summary evaluation sheet," which is actually a set of pages that condense the findings of the

⁴Two questions that might be candidate "fatal flaw" items, when responses to them were unacceptable (i.e., "no"), are the following from the validity section:

- Generally, the estimates of benefits, harms, and costs are consistent with the evidence presented in the guidelines document. (Yes, completely; yes, partially; or no) (Question 28)
- *Each major recommendation* is consistent with the estimated benefits, harms, and costs of the service or intervention (and thus with the strength of the evidence). (Yes, completely; yes, partially; or no) (Question 31).

assessment from the primary questions in the instrument. It is filled out only *after* the full instrument has been completed.

In the present version, "better" answers are recorded to the left of the response column, "worse" answers to the right. Thus, a quick scan of this sheet may provide an overall sense of the quality of the guidelines. Put another way, a clinician or other user ought to be able to apply the guideline effectively if all dimensions of the seven attributes (or, at a minimum, all seven attributes) are judged to be "satisfactory" and all omissions of information are considered "unimportant," as those terms were defined earlier. In this (ideal) situation, all notations on the summary sheet would be on the far left. By contrast, if many or most notations are on the right side of the response column, the user might wish to employ the guidelines only selectively or to request clarifications or revisions.

Several reviewers noted that completing the summary evaluation sheet is essentially a clerical task, provided the main part of the assessment instrument has been legibly and fully completed by one or more experts (as discussed earlier). The committee agrees and thus suggests that junior or clerical staff be given this responsibility, and that the instructions on the form indicate that users might wish to do so. Alternatively, the instrument or at least the recording of responses to its questions) might be computerized. In that case, the summary sheet could be an automatic product of the computer program. The value in pursuing more fully the possibilities of computerization of this form might be considerable.

FINAL CAUTIONS AND CAVEATS

The Ideal: Enemy of the Good

Some commentators noted the IOM's recognition in its 1990 report that most (if not all) guidelines in existence today would "fail" to meet the ideal of this instrument. Reviewers were concerned that "prematurely imposing excessive rigor" would discourage some guideline developers.⁵

⁵Comments from the American Nurses Association were particularly to the point: the criteria used to judge the adequacy of guidelines establishes a high standard that likely would seldom be achieved in reality. Several questions need careful consideration prior to accepting the criteria in this instrument:

1. Do the criteria . . . create false expectations of quality which is not achievable with current fiscal restrictions in health care?
 2. What are the potential legal and regulatory ramifications of accepting these criteria as representative of quality practice?
 3. How would these criteria eventually influence the costs of care through the pursuit of considerable evidence regarding the "best" method of treatment?
 4. Are these guidelines intended to weed out bad practice. or is the intention to demonstrate the "best," often misinterpreted as the "only" way of practicing?"
- (K. S. O'Connor, Division of Nursing Practice and Economics, American Nurses

At least one reviewer warned that the assessment process should not be used as a "second level expert panel" and cautioned that designing the assessment instrument process itself would take some care. The IOM committee agrees and, in that light, emphasizes that the educational uses of the instrument are more important than its assessment applications (in the near term at least).

Users of the Assessment Instrument

Busy practicing physicians or other clinicians are *not* the intended or anticipated appliers of this assessment instrument. Neither are policymakers, patients, or other nonprofessionals, although all may have some interest in the results. Assessors are expected to have, individually or collectively, expertise in three areas: clinical experience with patient populations covered by the guideline, research experience about the conditions or technologies covered by the guideline, and methodological expertise with techniques and processes of guideline development.

Because no one individual is likely to possess all three kinds of expertise, experience with the instrument may suggest that a "group," "panel," or "study section" approach will be needed to apply it satisfactorily. In this way, different individuals would be responsible for different parts of the assessment (particularly to determine validity). Furthermore, turning the full assessment into a review or evaluation that would be understandable to patients, practitioners, or policymakers will be a separate step, as noted earlier.

This provisional instrument thus proceeds on several assumptions. First, assessors (individually or collectively) are sufficiently schooled either in the methodologic issues inherent in guidelines development or in the clinical issues related to the main topic of the guidelines document that they are able to complete the bulk of the assessment instrument unaided. Second, questions about clinical topics or methods can be referred to appropriate experts when necessary and without undue delay. This kind of referral is particularly important if the AHCPR or some other entity acquires a specific mandate to certify or ratify guidelines from whatever source. Third, in some cases, having a dual or parallel (i.e., simultaneous) review may be a desirable tactic. Fourth, junior staff may well be used to assemble relevant material, perhaps to do an initial check of the document itself, perhaps to evaluate the document for the attribute of clarity, and to complete the summary evaluation sheet. Finally, the experts assembled or asked to apply the instrument (in its current form or any future, modified form) will be carefully trained in its use.

Availability of Supporting Material

Guidelines documents would become unmanageably long and unworkable for busy clinicians if all the information leading to specific recommendations were included in the

Association. in a letter to Marilyn Field, IOM study director, dated May 20, 1991.)

guideline itself. Nevertheless, the availability of such information somewhere is important. Meeting this expectation presents a difficult conflict for guideline developers, and it seriously complicates the task of assessing guidelines.⁶

At this time, the committee takes the position that as much information as possible should be synthesized into a guideline document, even if the formal guideline made widely available to practitioners and clinicians is a streamlined version. The assessment process is then to be directed at the complete document (with whatever supporting materials may be submitted with it), not the clinician's version. This stance accords with the committee's general goal of assessing the underlying quality of the guideline itself, not just the process by which it was developed.

For any of the uses to which the present instrument is put, the committee thus assumes that relevant documentation will be in the assessors' hands. This assumption is particularly important when the guidelines to be assessed have been developed by others, and especially if those organizations have approached AHCPR or another certifying entity with a specific request for ratification of the guidelines. Hence this instrument assumes that the pertinent information concerning the guidelines document—including information related to the process of development itself—is available for any review effort, with no provision for "later" or "on request" submission of information.

A consequence of this assumption is that this instrument is directed at "guideline documents," however those documents might be construed by the developers. Some guidelines may be contained within a single report, monograph, or other publication. Other guidelines may incorporate related publications by reference, particularly when developers have used a standardized methodology that is described elsewhere. Primary and secondary publications, reports, and records relating to the development of the guideline document being assessed should be assembled before the assessment exercise begins. This might include, for example, reviews and syntheses of the relevant scientific literature, but such a requirement would *not* extend to individual research reports and articles themselves.

Nevertheless, the committee recognizes that published guidelines may be "incomplete" because of limitations placed on the authors by editors and publishers (e.g., space constraints) and that important documentation may not be present. Therefore, if the instrument is applied to published guidelines developed by a group that does not deliberately seek to have its guidelines so assessed (and

⁶For example, the AHCPR Forum panel that has been working on the issue of managing depression in community-based settings started originally with 50,000 citations to the literature, reviewed between 4,000 and 6,000 articles, and based its guidelines document on about 400 relevant articles (J. J. Strain, member of the IOM committee, in a memorandum to Kathleen Lohr and Marilyn Field, dated June 17, 1991). There is no possibility that assessors of the guideline document could replicate that experience or even undertake to review the final set of relevant articles.

volunteer the supporting material as assumed above) additional material may need to be gathered from those authors in order to apply this instrument fairly.

Standardized Format Versus Narrative Evaluation

Regardless of what approach to assessing guidelines is finally adopted and what level of expertise the evaluators possess, several basic complexities must be acknowledged. For instance, simply assessing whether guideline developers explain or document a certain piece of information does not allow one to discriminate a comprehensive disclosure from one of poor quality. Similarly, lack of disclosure may have a significant or only, trivial impact on the clinical usefulness and validity of a guideline. No structured instrument of practical length is likely to be able to accommodate these nuances across guideline documents of many different types. Thus some narrative, global assessment may always be desirable, if not absolutely necessary, if assessments of guidelines are to be useful for a wide set of audiences.

In developing the instrument, the committee asked the first set of reviewers to comment on a "handbook" approach as an alternative to the formal instrument. This approach would provide guidelines assessors with some instructions about the attributes of guidelines to be evaluated and would require them to prepare a narrative evaluation statement, but it would not produce specific responses to specific questions.

Some respondents preferred the "objective, criterion-based" review (i.e. the formal instrument), noting that it might yield "a more standardized evaluation strategy" and "potential benefits such as greater efficiency and reliability, a more readily digested assessment of the strengths and weaknesses of a guideline, and the ability to draw more 'objective' comparisons among a collection of guidelines." Given that no clear preference for the handbook approach emerged, the committee did not pursue this approach further. However, the desire expressed by several reviewers for a narrative, summary statement about a guideline document probably reflects some discomfort with an assessment strategy based solely on the question-and-answer format of the present instrument.

Further Pretesting and Experience with the Instrument

In developing this instrument, the committee recognized the need for more practical experience with it. The present version incorporates revisions suggested by the large number of reviewers, mainly from medical specialty societies, who critiqued an earlier version and in some cases applied it to actual guidelines, as well as changes pursuant to the IOM/NRC review. Nevertheless, the committee takes the view that further application and pretesting of this provisional form should be conducted.

That testing should determine answers to the following questions: Is the instrument too long and too complicated for practical routine use? Does experience applying the instrument as an assessment tool make it easier to use, as several reviewers believed it might? Can shortcuts be found in applying it? For instance, is it useful to have junior staff

make an initial check to determine whether all relevant materials appear to be in the guidelines document package or to make a first-assessment pass through the guidelines document itself? Are the results of the assessment consistent with results of any pretests or early evaluations of the guidelines in actual practice?

The present committee takes no stand on how extensive such pre- or pilot-testing might be—for instance, on the number of guidelines that should be assessed to determine the reliability, validity, and practicality of the current form. Two factors are relevant. First, the committee had neither the time nor the resources to pursue these issues further (and certainly not to carry out such activities itself). Second, it considers that such testing might need to be specific to the potential user groups and that setting *a priori* rules risks making them too rigorous or too confining for all purposes.

THE PROVISIONAL ASSESSMENT INSTRUMENT

The form reproduced in the last part of this appendix has three main parts. First is a *general information sheet*, with space for the following items to be briefly described: clinical diagnoses or conditions; health practices, services, or technologies; target populations; primary settings of care; primary types of clinicians targeted; stated purposes of the guideline; source, author, or developer of the guideline document; person to contact for further information about the guideline document; date of issue of the guideline document; and name/affiliation of assessor(s). The second section is the *full instrument* itself, with self-contained instructions. The third section is the *summary evaluation sheet*.

ACKNOWLEDGMENTS

The idea of creating an instrument by which to assess the soundness of clinical practice guidelines can be traced through the work of three Institute of Medicine committees: the Committee to Design a Strategy for Quality Review and Assurance in Medicare, the Committee to Advise the Public Health Service on Clinical Practice Guidelines, and the present Committee on Clinical Practice Guidelines. The present committee wishes, therefore, to acknowledge the groundbreaking efforts of its predecessor panels in this rapidly evolving arena.

Production of this provisional instrument would not have been possible without the assistance of many individuals and organizations. We are indebted to the members of various medical specialty societies who reviewed and in some cases voluntarily applied and tested an earlier draft instrument; we especially thank Betty King executive director of IMCARE (Internal Medicine Center to Advance Research and Education) for her efforts in organizing a broad review of the instrument by the IMCARE task force on practice guidelines. We thank Anne-Marie Audet of the Health Institute, New England Medical Center Hospitals, for her enthusiastic efforts with the first draft of the instrument.

The project under which this instrument was developed was supported by The John A. Hartford Foundation and by the Agency for Health Care Policy and Research, U.S. Department of Health and Human Services, under Contract No. 282-90-0018. The views

presented are those of the Institute of Medicine Committee on Clinical Practice Guidelines and the authors, and are not necessarily those of The funding organizations.

Finally, we extend our appreciation to our Institute of Medicine (IOM) colleagues on this project, Molla Donaldson and Holly Dawkins; to the project's senior secretaries, Theresa Nally and Donna Thompson; and to other members of the IOM staff who provided timely and helpful comments, including Christopher Howson, Michael Stoto, and Malin VanAntwerp.

REFERENCES

AMA (American Medical Association). "Evaluation of a Recently Completed Practice Parameter." Chicago: American Medical Association, 1990.

IOM (Institute of Medicine). Clinical Practice Guidelines: Directions for a New Program M.J. Field and K.N. Lohr, editors. Washington, DC: National Academy Press. 1990.

ASSESSMENT INSTRUMENT

PART ONE. GENERAL INFORMATION SHEET

TO THE ASSESSOR: Please complete this sheet with brief statements about the content of the guideline document you are reviewing. Use whatever information can be found in the document. If you cannot find the relevant information or are uncertain about the appropriate response, indicate “not specified” or “uncertain.”

TITLE OF GUIDELINE DOCUMENT _____

1. Clinical diagnoses or conditions
2. Main health practices, services, or technologies considered
3. Target populations (e.g., age, sex, income level, health status)
4. Primary settings of care (e.g., primary or specialty; nursing home)
5. Primary types of clinicians targeted (e.g., profession; specialty)
6. Stated purposes, aims, or goals of the guideline document
7. Source, author, or developer of guideline document
8. Individual to contact for further information about the guideline document (name, organization, phone number)
9. Date of issue of the guideline document
10. Name/affiliation of assessor(s)

PART TWO. ASSESSMENT INSTRUMENT

Background

This instrument itself has seven sections, each corresponding to one of the seven attributes of guidelines to be evaluated. Each section begins with a brief definition of the attribute and then is divided into segments that deal with important dimensions of that attribute.

Instructions: Illustrative Example

Each segment begins with a descriptive question that you should answer yes or no. The responses will then direct you to move to a specific next question. Space is provided for "Comments" throughout the instrument.

For example, in the section on clinical applicability, the first question reads (in part) as shown below, and you are instructed to check "yes" or "no" and then answer the appropriate subquestion:

1. THE GUIDELINE DOCUMENT DESCRIBES THE PATIENT POPULATIONS TO WHICH THE GUIDELINES ARE MEANT TO APPLY.

☐ Yes (Go to 1.1) ☐ No (Go to 1.2)

1.1. THE DESCRIPTION OF THE PATIENT POPULATIONS IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

>> GO TO QUESTION 2 >>

1.2. OMISSION OF A DESCRIPTION OF THE PATIENT POPULATIONS IS:

☐ Unimportant omission ☐ Minor omission ☐ Major omission

In some circumstances you are asked to judge a set of variables---specific elements to consider in evaluating an attribute---and to arrive at a "global" answer to a particular

question. When this occurs, you should (1) start with the set of items that are identified alphanumerically (e.g., 1a, 1b, . . .) and answer them directly and then (2) combine those answers into a summary evaluation to determine whether the guideline document has dealt with that particular issue in a satisfactory, conditionally satisfactory, or unsatisfactory manner.

You should then go on to the next question or to the next section, as directed. In the absence of a specific direction, go to the very next question.

Definitions of Terms

The questions in this instrument ask for three different types of responses. The meaning of these response categories is as follows:

Yes and no. Most of the main questions concern the presence of a discussion or piece of information about a particular attribute. Generally, the "yes" and "no" responses direct you to answer follow-up questions. For these items, response choices are "satisfactory," "conditionally satisfactory," and "unsatisfactory," or "unimportant omission," "minor omission," and "major omission." These terms are further defined below.

Satisfactory. You can judge information about a particular attribute as satisfactory if all critical elements have been considered and presented. For example, the discussion or description should be thorough and comprehensive; the guideline developers should have based their work on appropriate and correct information; and they should have used appropriate methods.

Conditionally satisfactory. The description or discussion of an attribute or dimension is conditionally satisfactory if some, but not all, of the critical elements have been considered and presented. For example, the discussion of a particular aspect of the attribute may be vague or incomplete; alternatively, the guidelines developers may have disregarded important information in reaching their recommendations or improperly used certain kinds of methods. These problems with the guideline document may not prevent a clinician from using it or understanding its recommendations, but they may affect its overall usefulness or call certain recommendations into question. Revisions would be presumed to improve the guidelines, but they would not be mandatory or essential.

Unsatisfactory. You can determine the description or discussion about a specific attribute to be unsatisfactory if most of the critical elements have not been considered or presented. For example, well-known pieces of clinical information (or the views of multiple specialists with an interest in the guidelines topic) may have been ignored; methods of analysis may have been misapplied; or recommendations may be based on faulty information poor logic, or both. In such a case, it would be difficult if not impossible to judge the quality of the process of guideline development or the soundness of the resulting guidelines and recommendations, or both; certainly they could not be assessed as satisfactory, and serious thought must be given to augmenting or revising the guideline document before it is promoted further.

Unimportant omissions. You can regard the omission of a description or discussion about a specific attribute or dimension of an attribute as unimportant if it (1) is likely to have no meaningful impact on the ability of a guideline user, such as a practitioner or patient, to apply the guideline effectively in the clinical decision-making process and (2) does not prevent you from easily and independently assessing that aspect of the guideline document.

Minor omissions. The omission of a description or discussion about a specific attribute or dimension is of minor importance if it (1) is likely to have only a little negative impact on the ability of a guideline user to apply the guideline effectively in the clinical decision-making process and (2) does not prevent you from assessing that aspect of the guideline document.

Major omissions. You can determine the omission of a description or discussion about a specific attribute or dimension to be a major problem if it (1) is likely to prevent a guideline user from applying the guideline effectively in the clinical decision-making process or (2) prevents you from making an independent assessment about that aspect of the guideline document.

Not applicable or don't know. In some situations, the question may not be applicable to the guideline document you are evaluating. When that occurs, simply mark "NA" for "not applicable" or "DK" for "don't know."

Comments. In other situations, you may find it difficult to arrive at a single answer to the question, especially if the guideline document you are evaluating is very complex or if necessary background information appears to be missing. In these cases, you can record additional remarks or qualifying statements about your response in the "Comments" sections.

Finally, if you believe that the guideline document is so complex, clinically esoteric, or methodologically sophisticated that it warrants additional, outside expert review, please note your comments at the end of the full instrument.

I. CLINICAL APPLICABILITY

Clinical applicability, or the scope of the guideline, means three things in the context of this instrument. First, guidelines should be written to cover as inclusive a patient population as possible, consistent with knowledge about critical clinical and sociodemographic factors relevant for the condition or technology in question. To that end, the patient population(s) covered should be described as accurately and precisely as possible. Second, if patient populations that might be expected to be covered by the guideline are not, then the document discusses why those populations have been excluded; that is, it identifies the patient populations the guidelines are not meant to serve or apply to. Third, when the clinical conditions or problems covered by the guideline are likely to be complex, or when the guideline recommendations may be contingent on complex patterns of clinical factors, those points should be explicitly covered in the guideline document.

This attribute requires that two things be true about the guideline document. First, the guideline document accurately and precisely states how broad or narrow the patient population(s) are to which the guidelines are meant to apply, describes the actual population(s) to which statements apply, and describes the population(s) to which statements are not meant to apply. Population(s) may be described in terms of diagnosis, pathophysiology, severity of primary disease, presence of coexisting diseases, age, sex, race, social support systems, and other characteristics. Second, it notes and discusses any complex clinical issues that may arise for this patient population.

1. THE GUIDELINE DOCUMENT DESCRIBES THE PATIENT POPULATIONS TO WHICH THE GUIDELINES ARE MEANT TO APPLY.

☐ Yes (Go to Question 1.1) ☐ No (Go to Question 1.2)

1.1. THE DESCRIPTION OF THE PATIENT POPULATIONS IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

>> GO TO QUESTION 2 >>

1.2. OMISSION OF THE DESCRIPTION OF THE PATIENT POPULATION(S) IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

2. THE GUIDELINE DOCUMENT DISCUSSES COMPLEX CLINICAL PROBLEMS THAT CAN BE EXPECTED FOR THE POPULATION(S) COVERED BY THE GUIDELINES.

☐ Yes (Go to Question 2.1)

☐ No (Go to Question 2.2)

☐ Not Applicable (Go to Question 3)

2.1. THE DISCUSSION OF EXPECTED COMPLEX CLINICAL PROBLEMS IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

>> GO TO QUESTION 3 >>

2.2. OMISSION OF THE DISCUSSION OF EXPECTED COMPLEX CLINICAL PROBLEMS IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

3. THE GUIDELINE DOCUMENT GIVES A RATIONALE FOR EXCLUDING PATIENT POPULATION(S).

☐ Yes (Go to Question 3.1) ☐ No (Go to Question 3.2)

3.1. THE RATIONALE FOR EXCLUDING CERTAIN PATIENT POPULATION(S) IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

>> GO TO II. CLINICAL FLEXIBILITY >>

3.2. OMISSION OF THE RATIONALE FOR EXCLUDING CERTAIN PATIENT POPULATION(S) IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

II. CLINICAL FLEXIBILITY

Clinical flexibility means that two mediating factors should be addressed, in the guideline document. First, it should identify major foreseeable exceptions to or options for applying the guidelines, if any exist. Second, it should discuss the role of patient preferences for different courses of health care for those conditions or technologies in which patient values and preferences may be important decision-making factors (for example, being able to choose in an informed way between surgery and watchful waiting).

This attribute requires the guideline document to discuss two topics. First are situations (if any) in which socially relevant factors permit an exception to be made in applying the guidelines. These factors could include the home and family situation of the patient, clinical constraints on the health care delivery setting (e.g., no intensive care beds, no 24-hour anesthesiologist), *nonclinical* constraints on the health care delivery setting (e.g., inadequate information systems), or all of these; if no such factors exist, the guideline document should say so. Second is the role of patient preferences for different possible outcomes of care, when the appropriateness of a clinical intervention involves a substantial element of personal choice or values on the part of the patient. For example, this discussion may include information as to major points on which preferences may diverge for the case in hand, specific points to consider in eliciting patient preferences, and means of integrating patient views in the decisionmaking process.

4. THE GUIDELINE DOCUMENT PROVIDES SPECIFIC INFORMATION ABOUT SITUATIONS IN WHICH CLINICAL EXCEPTIONS MIGHT BE MADE IN APPLYING THE GUIDELINES.

☐ Yes, the document gives information about clinical exceptions
(Go to Question 4.1)

☐ No, the document says nothing about clinical exceptions
(Go to Question 4.2)

4.1. THE INFORMATION OR STATEMENT ABOUT *CLINICAL* EXCEPTION IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

>> GO TO QUESTION 5 >>

4.2. OMISSION OF INFORMATION OR A STATEMENT ABOUT *CLINICAL* EXCEPTIONS IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

5. THE GUIDELINE DOCUMENT PROVIDES SPECIFIC INFORMATION ABOUT *NONCLINICAL* SITUATIONS IN WHICH EXCEPTIONS MIGHT BE MADE IN APPLYING THE GUIDELINES.

☐ Yes, the document gives information about nonclinical exceptions
(Go to Question 5.1)

☐ No, the document says nothing about nonclinical exceptions
(Go to Question 5.2)

5.1. THE INFORMATION OR STATEMENT ABOUT *NONCLINICAL* EXCEPTIONS IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

>> GO TO QUESTION 6 >>

5.2. OMISSION OF INFORMATION OR A STATEMENT ABOUT *NONCLINICAL* EXCEPTIONS IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

6. THE GUIDELINE DOCUMENT DISCUSSES THE ROLE OF PATIENT PREFERENCES, AS THEY RELATE TO HEALTH CARE DECISIONS IN THE PARTICULAR CASE THAT THE GUIDELINES COVER.

☐ Yes (Go to Question 6.1) ☐ No (Go to Question 6.2)

6.1. THE DISCUSSION OF PATIENT PREFERENCES IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

>> GO TO QUESTION 7 >>

6.2. OMISSION OF DISCUSSION OF PATIENT PREFERENCES IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

>> GO TO III. RELIABILITY/REPRODUCIBILITY >>

7. THE GUIDELINE DOCUMENT DESCRIBES HOW PATIENT PREFERENCES WERE TAKEN INTO ACCOUNT DURING THE GUIDELINE DEVELOPMENT PROCESS.

☐ Yes (Go to Question 7.1) ☐ No (Go to Question 7.2)

**7.1. THE DISCUSSION OF HOW PATIENT PREFERENCES WERE CONSIDERED
IN DEVELOPING THE GUIDELINE IS:**

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

>> GO TO III. RELIABILITY/REPRODUCIBILITY >>

**7.2 .OMISSION OF THE DISCUSSION OF PATIENT PREFERENCES IN
DEVELOPING THE GUIDELINE IS:**

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

III. RELIABILITY/REPRODUCIBILITY

Reliability and reproducibility for the purpose of assessing guidelines means that, given the same circumstances, essentially the same set of guidelines would be developed by a second group; further, the terms mean that, ideally, the guidelines are or would be interpreted and applied Consistently by practitioners or other appropriate parties.

Reliability and reproducibility of a guideline document is not likely ever to be assessable empirically. To approach these concepts, therefore, this attribute requires either that guidelines be subjected to some form of explicit, independent review by a group (or groups) other than the original developers, where that group (or groups) is equivalent in expertise and other factors to the original developers, or that the guideline recommendations have been pretested in some manner, or both. (Pretesting can be done in actual delivery settings or on prototypical cases.) If no such review or pretesting has been done, then the guidelines must explain the reasons.

8. THE GUIDELINES WERE SUBJECTED TO INDEPENDENT REVIEW BY EXPERTS OR OUTSIDE PANELS.

☐ Yes (Go to Question 8.1) ☐ No (Go to Question 9)

8.1. THE DISCUSSION OF INDEPENDENT REVIEW IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

>> GO TO QUESTION 10 >>

9. THE GUIDELINE DOCUMENT EXPLAINS THE LACK OF INDEPENDENT REVIEW.

☐ Yes (Go to Question 9.1) ☐ No (Go to Question 9.2)

9.1. THE EXPLANATION OF THE LACK OF INDEPENDENT REVIEW IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

>> GO TO QUESTION 10 >>

9.2. OMISSION OF AN EXPLANATION OF THE LACK OF INDEPENDENT REVIEW IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

10. THE GUIDELINES WERE PRETESTED IN SOME MANNER.

☐ Yes (Go to Question 10.1) ☐ No (Go to Question 11)

10.1. THE DISCUSSION OF PRETESTING IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

>> GO TO IV. VALIDITY >>

11. THE GUIDELINE DOCUMENT EXPLAINS THE LACK OF PRETESTING.

☐ Yes (Go to Question 11.1) ☐ No (Go to Question 11.2)

11.1. THE EXPLANATION OF THE LACK OF PRETESTING IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

>> GO TO IV. VALIDITY >>

11.2. OMISSION OF AN EXPLANATION OF THE LACK OF PRETESTING IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

IV. VALIDITY: DEFINITION AND EVALUATION QUESTIONS

Validity of practice guidelines means, conceptually, that if they are followed, then they will lead to the health and cost outcomes projected for them. Validity must be judged primarily by reference to the substance and quality of the evidence cited, the means used to evaluate the evidence, and the relationship between the evidence and the recommendations. Validity is the most critical attribute and the most difficult to assess. **Although this section contains 22 questions, questions 28 and 31 are, together, of special importance because they constitute an overall evaluation of this attribute.**

This attribute requires that five things be true for the guideline document. First, the collection, synthesis, and interpretation of scientific evidence must be documented and of satisfactory quality; ideally, each major recommendation will be described as based on "excellent," "acceptable," or "weak" evidence, or with a similar set of descriptive terms.

Second, both qualitative and quantitative statements about health benefits and harms/risks appear in the guideline document, and insofar as possible those estimates are tied to and justified by the evidence amassed as part of the literature review and analysis. For example, a qualitative statement about benefits might read "screening mammography should lead to a decrease in breast cancer mortality"; a similar statement about harms and risks might read "screening mammography can lead to false-positive results and to unnecessary work-up and anxiety." Quantitative statements might read, respectively, "screening mammography in women 50 years of age may reduce mortality from 20 percent to 60 percent" and "among one million women 40 to 50 years of age, radiation from 10 mammography examinations can be expected to cause about 60 new breast cancers." In all cases, such statements should be based on evidentiary information insofar as possible, and appropriate qualifiers or caveats noted when the evidence is weak or conflicting or when the estimates are based on consensus techniques such as expert panels or group judgment methods.

Third, both qualitative and quantitative statements about expected health costs or expenditures appear in the guideline document; the same requirements about the link between the guideline estimates and the data sources should be met, and the same degree of specificity about patient groups should be observed. In addition, the document should be clear as to whether costs referred to are the total for the patient group or the per-patient figure. For example, "use of laparoscopic techniques to treat cholecystitis should reduce the direct and indirect costs associated with using cholecystectomy as the main patient management approach" might be a suitable qualitative statement concerning costs, and "use of laparoscopic techniques in the treatment of cholecystitis may reduce the costs of treatment as much as 75 percent by the end of the decade by reducing hospitalization and time for post-operative (i.e., post-cholecystectomy) morbidity and recovery" might be an appropriate quantitative statement about estimated costs.

Fourth, specific recommendations are clearly tied to and justified by the estimated benefits, harms, and costs provided within the document.

Fifth, conflicts between this set of guidelines and any other independent sets (and their respective recommendations), if any, must be explicitly discussed.

Strength of Scientific Evidence and Professional Consensus

12. THE GUIDELINE DOCUMENT SPECIFICALLY DESCRIBES THE METHOD(S) USED TO COLLECT (I.E., IDENTIFY AND RETRIEVE) THE SCIENTIFIC EVIDENCE ON WHICH RECOMMENDATIONS ARE BASED.

☐ Yes (Go to Question 12.1) ☐ No (Go to Question 12.2)

12.1. ASSESSOR: Respond to Items 12a-d, below, to assess the methods for collecting scientific evidence; then answer Question 12.1, using your best judgment as to the overall rating for this element of validity. Other factors you judge important should be specifically recorded under "Comments or Other Factors."

12a. The criteria used to include and/or exclude studies are:

☐ Adequate ☐ Inadequate ☐ Not given/described

12b. The search strategy is:

☐ Adequate ☐ Inadequate ☐ Not given/described

12c. The sources of information are:

☐ Adequate ☐ Inadequate ☐ Not given/described

12d. Major studies or other sources of information have been identified.

☐ Yes ☐ No (Specify) ☐ Don't know

Now answer:

12.1. THE METHOD(S) OF COLLECTING SCIENTIFIC EVIDENCE IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments or Other Factors:

>> GO TO QUESTION 13 >>

12.2. THE LACK OF A CLEAR METHOD FOR COLLECTING THE SCIENTIFIC EVIDENCE IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

13. THE GUIDELINE DOCUMENT GIVES ADEQUATE REFERENCES OR CITATIONS TO THE SOURCES OF INFORMATION USED IN DEVELOPING THE GUIDELINES.

☐ Yes ☐ No

Comments:

14. THE GUIDELINE DOCUMENT DISCUSSES IN GENERAL TERMS THE STRENGTH OF THE SCIENTIFIC EVIDENCE ON WHICH RECOMMENDATIONS ARE BASED.

Comments:

15. THE GUIDELINE DOCUMENT EXPLICITLY RATES THE STRENGTH OF THE SCIENTIFIC EVIDENCE.

☐ Yes (Go to Question 15.1) ☐ No (Go to Question 15.2)

15.1 ASSESSOR: Respond to Items 15a-15f, below, to determine whether the method used to rate the strength of the scientific evidence is adequate; then answer Question 15.1 below, using your best judgment as to the overall rating for this element of validity. Other factors you judge important should be specifically recorded under "Comments or Other Factors."

15a. Characteristics of studies used as a basis for guidelines have been described.

☐ Yes ☐ No

15b. Strengths and weaknesses of studies used as a basis for guidelines have been noted.

☐ Yes ☐ No

15c. The way the characteristics, strengths, and weaknesses of studies used as a basis for guidelines have been taken into account (for instance, an explicit weighting scheme) has been clearly described.

☐ Yes ☐ No

15d. The way the characteristics, strengths, and weaknesses of studies used as a basis for guidelines have been taken into account (for instance, an explicit weighting scheme) is:

☐ Adequate ☐ Inadequate ☐ Not given/described

15e. The discussion in the document of possible threats to internal validity and reliability of studies included in the scientific evidence supporting the guidelines is:

☐ Adequate ☐ Inadequate ☐ No discussion given

15f. The discussion in the document of possible threats to external validity and generalizability of studies included in the scientific evidence supporting the guidelines is:

☐ Adequate ☐ Inadequate ☐ No discussion given

Now answer:

15.1. OVERALL, THE METHOD USED TO RATE OR WEIGHT THE SCIENTIFIC EVIDENCE IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments or Other Factors:

>> GO TO QUESTION 16>

15.2. THE LACK OF ANY GENERAL DISCUSSION OR EXPLICIT RATING OF THE STRENGTH OF THE SCIENTIFIC EVIDENCE IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

16. IF A FORMAL METHOD OF SYNTHESIS IS USED TO COMBINE THE SCIENTIFIC EVIDENCE QUANTITATIVELY OR OTHERWISE TO DEVELOP SUMMARY OUTCOME MEASURES THAT REFLECT THE STRENGTH OF THE SCIENTIFIC EVIDENCE, THEN THE GUIDELINE DOCUMENT EXPLICITLY DESCRIBES THE METHOD.

☐ Yes, method used and described (Go to Question 16.1)

☐ No, method used but not described (Go to Question 16.2)

☐ No, no formal method of synthesis used (Go to Question 18)

16.1. ASSESSOR: Respond to Items 16a-16c, below, to determine whether formal methods for synthesizing scientific evidence are satisfactory; then answer Question 16.1 below, using your best judgment as to the overall rating for this element of validity. Other factors you judge important should be specifically recorded under "Comments or Other Factors."

16a. The meta-analytic method(s) is:

☐ Adequate ☐ Inadequate ☐ Not applicable/used

16b. The decision-analytic model(s) is:

☐ Adequate ☐ Inadequate ☐ Not applicable/used

16c. Other systematic information synthesis method(s) is:

☐ Adequate ☐ Inadequate ☐ Not applicable/used

Now answer:

16.1 OVERALL, THE FORMAL METHOD(S) USED TO SYNTHESIZE OR COMBINE SCIENTIFIC EVIDENCE IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments or Other Factors:

>> GO TO QUESTION 17 >>

16.2. OMISSION OF A DESCRIPTION OF THE METHOD(S) OF SYNTHESIZING THE SCIENTIFIC EVIDENCE IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

17. GIVEN THAT A FORMAL METHOD OF SYNTHESIS IS USED TO COMBINE THE SCIENTIFIC EVIDENCE QUANTITATIVELY OR OTHERWISE TO DEVELOP SUMMARY OUTCOMES MEASURES, THE GUIDELINE DOCUMENT EXPLICITLY REPORTS THE RESULTS OF THAT SYNTHESIS.

☐ Yes, method used and results reported (Go to Question 17.1)

☐ No, method used but results not reported (Go to Question 17.2)

17.1. RESULTS OF INFORMATION SYNTHESIS ARE:

☐ Satisfactory (e.g., summary outcome measure(s) with confidence intervals or discussion of uncertainty)

_____Conditionally satisfactory (e.g., summary outcome measure(s) without confidence intervals or discussion of uncertainty)

___Unsatisfactory (e.g., outcome measure(s) are not interpretable, are inconsistent, or are otherwise questionable or erroneous). (Specify)

Comments:

>> GO TO QUESTION 18 >>

17.2. OMISSION OF RESULTS OF SYNTHESIS IS:

___Unimportant ___ Minor omission ___ Major omission

Comments:

18. IF FORMAL EXPERT OR GROUP JUDGMENT TECHNIQUES ARE USED TO REACH PROFESSIONAL CONSENSUS, THEN THE GUIDELINE DOCUMENT EXPLICITLY DESCRIBES THE TECHNIQUES.

___Yes, techniques used and described (Go to Question 18.1)

___No, techniques used but not described (Go to Question 18.2)

___No, no formal expert or group judgment techniques used (Go to Question 19)

18.1. THE EXPERT OR GROUP JUDGMENT TECHNIQUES ARE:

___Satisfactory ___Conditionally satisfactory ___Unsatisfactory (Specify)

Comments:

>> GO TO QUESTION 19 >>

18.2. OMISSION OF A DESCRIPTION OF THE EXPERT OR GROUP JUDGMENT TECHNIQUES IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

19. GIVEN THAT EXPERT OR GROUP JUDGMENT METHOD(S) ARE USED TO REACH PROFESSIONAL CONSENSUS, THE GUIDELINE DOCUMENT EXPLICITLY GIVES INFORMATION ABOUT THE STRENGTH OF PROFESSIONAL CONSENSUS.

☐ Yes, techniques used and information given (Go to Question 19.1)

☐ No, techniques used but information not given (Go to Question 19.2)

19.1. THE INFORMATION ABOUT THE STRENGTH OF PROFESSIONAL CONSENSUS IS:

☐ Satisfactory (e.g., levels of professional consensus given for all major points in the guidelines)

☐ Conditionally satisfactory (e.g., levels of professional consensus given for some, but not all, major points in the guidelines)

☐ Unsatisfactory (e.g., levels of professional consensus are not interpretable, are inconsistent, or are otherwise questionable or erroneous). (Specify)

Comments:

>> GO TO QUESTION 20 >>

19.2. OMISSION OF EXPLICIT INFORMATION ABOUT THE STRENGTH OF PROFESSIONAL CONSENSUS IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

20. THE GUIDELINE DOCUMENT PROVIDES A QUALITATIVE DESCRIPTION OF THE HEALTH BENEFITS THAT ARE EXPECTED FROM A SPECIFIC HEALTH PRACTICE.

☐ Yes (Go to Question 20.1) ☐ No (Go to Question 20.2)

20.1. THE QUALITATIVE DESCRIPTION OF HEALTH BENEFITS IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

>> GO TO QUESTION 21 >>

20.2. OMISSION OF A QUALITATIVE DESCRIPTION OF HEALTH BENEFITS IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

21. THE GUIDELINE DOCUMENT PROVIDES A QUALITATIVE DESCRIPTION OF THE POTENTIAL HARMS OR RISKS THAT MAY OCCUR AS A RESULT OF A SPECIFIC HEALTH PRACTICE.

☐ Yes (Go to Question 21.1) ☐ No (Go to Question 21.2)

Comments:

21.1. THE QUALITATIVE DESCRIPTION OF POTENTIAL HARMS OR RISKS IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

>> GO TO QUESTION 22 >>

21.2. OMISSION OF A QUALITATIVE DESCRIPTION OF POTENTIAL HARMS OR RISKS IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

Health Benefits and Harms/Risks: Quantitative Information

22. THE GUIDELINE DOCUMENT PROVIDES QUANTITATIVE INFORMATION OR ESTIMATES ABOUT THE HEALTH BENEFITS TO BE EXPECTED AS A RESULT OF A SPECIFIC HEALTH PRACTICE.

☐ Yes (Go to Question 22.1) ☐ No (Go to Question 22.2)

22.1. THE QUANTITATIVE INFORMATION ABOUT THE HEALTH BENEFITS IS:

___ Satisfactory (e.g., one or more measures of benefits, including accurate summary or composite measures, with confidence intervals or discussion of uncertainty)

___ Conditionally satisfactory (e.g., one or more measures of benefits, without confidence intervals or discussion of uncertainty)

___ Unsatisfactory (e.g., measures are not interpretable, are inconsistent, or are otherwise questionable or erroneous). (Specify)

Comments:

>> GO TO QUESTION 23 >>

22.2. OMISSION OF QUANTITATIVE INFORMATION AND ESTIMATION OF HEALTH BENEFITS IS:

___ Unimportant ___ Minor omission ___ Major omission

Comments:

>> GO TO QUESTION 24 >>

23. THE GUIDELINE DOCUMENT PROJECTS HEALTH BENEFITS OR OUTCOMES IN TERMS OF ADDITIONAL LIFE EXPECTANCY OR SIMILAR MEASURES, SUCH AS QUALITY-ADJUSTED LIFE YEARS.

___ Yes ___ No ___ Not applicable/not necessary

Comments:

24. THE GUIDELINE DOCUMENT PROVIDES QUANTITATIVE INFORMATION OR ESTIMATES ABOUT THE POTENTIAL HARMS OR RISKS OCCURRING AS A RESULT OF A SPECIFIC HEALTH PRACTICE.

☐ Yes (Go to Question 20.1) ☐ No (Go to Question 20.2)

24.1. THE QUANTITATIVE INFORMATION ABOUT POTENTIAL HARMS OR RISKS OCCURRING AS A RESULT OF A SPECIFIC HEALTH PRACTICE IS:

☐ Satisfactory (e.g., one or more measures of harms or risks, including summary or composite measures, with confidence intervals or discussion of uncertainty)

☐ Conditionally satisfactory (e.g., one or more measures of harms or risks, without confidence intervals or discussion of uncertainty)

☐ Unsatisfactory (e.g., measure(s) are not interpretable, are inconsistent, or are otherwise questionable or erroneous). (Specify)

Comments:

>> GO TO QUESTION 25 >>

24.2. OMISSION OF QUANTITATIVE INFORMATION ABOUT POTENTIAL HARMS OR RISKS IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

Health Costs: Qualitative Description

25. THE GUIDELINE DOCUMENT PROVIDES A *QUALITATIVE DESCRIPTION OF THE HEALTH COSTS* OR EXPENDITURES THAT ARE EXPECTED FROM A SPECIFIC HEALTH PRACTICE

☐ Yes (Go to Question 25.1) ☐ No (Go to Question 25.2)

25.1. THE QUALITATIVE DESCRIPTION OF EXPECTED HEALTH COSTS OR EXPENDITURES IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

25.2. OMISSION OF A QUALITATIVE DESCRIPTION OF EXPECTED HEALTH COSTS OR EXPENDITURES IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

Health Costs: Quantitative Description

26. THE GUIDELINE DOCUMENT PROVIDES QUANTITATIVE INFORMATION OR ESTIMATES ABOUT THE HEALTH COSTS OR EXPENDITURES THAT ARE EXPECTED AS A RESULT OF A SPECIFIC HEALTH PRACTICE.

☐ Yes (Go to Question 26.1) ☐ No (Go to Question 26.2)

26.1. ASSESSOR: Respond to Items 26a-26e, below, to determine whether potential costs and expenditures have been estimated in a satisfactory manner; then answer Question 26.1, using your best judgment as to the overall rating for this

element of validity. Other factors you judge important should be specifically recorded under "Comments or Other Factors."

26a. The cost estimates are done for major subgroups of the patient population, e.g., major risk groups, and for major clinical (diagnostic, therapeutic, etc.) alternatives.

☐ Yes ☐ No

26b. The cost estimates include all the services necessary to achieve the health benefits that are assumed to be achievable.

☐ Yes ☐ No

26c. The cost estimates specify number(s) of services that may be added, substituted, and/or eliminated if the guideline recommendations are followed.

☐ Yes ☐ No

26d. The cost estimates specify charges, production costs, or similar information for the services that may be added, substituted, and/ or eliminated if the guideline recommendations are followed.

☐ Yes ☐ No

26e. The quantitative method(s) used to estimate costs is

☐ Appropriate ☐ Inappropriate

Now answer:

26.1 THE QUANTITATIVE INFORMATION ABOUT EXPECTED HEALTH COSTS OR EXPENDITURES IS:

☐ Satisfactory (e.g., one or more estimates of costs, including accurate summary or composite measures, with ranges of uncertainty)

☐ Conditionally satisfactory (e.g., one or more estimates of Costs without ranges of uncertainty)

☐ Unsatisfactory (e.g., cost estimates are not interpretable, are inconsistent, or are otherwise questionable or erroneous). (Specify)

Comments or Other Factors:

>> GO TO QUESTION 27 >>

26.2. OMISSION OF QUANTITATIVE INFORMATION ABOUT EXPECTED HEALTH COSTS OR EXPENDITURES IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

>> GO TO QUESTION 28 >>

27. IF HEALTH BENEFITS ARE PROJECTED IN TERMS OF ADDITIONAL LIFE EXPECTANCY OR SIMILAR MEASURES, SUCH AS QUALITY-ADJUSTED LIFE YEARS, THEN THE COST PER UNIT OF EACH IDENTIFIED BENEFIT IS ESTIMATED.

☐ Yes, benefits projected in such terms and cost per unit estimated
☐ No, benefits projected in such terms but cost per unit not estimated
☐ Not applicable, benefits not so projected and cost per unit not estimated

28. GENERALLY, THE ESTIMATES OF BENEFITS, HARMS, AND COSTS ARE CONSISTENT WITH THE STRENGTH OF THE EVIDENCE PRESENTED IN THE GUIDELINE DOCUMENT.

☐ Yes, completely ☐ Yes, partially ☐ No

Comments:

29. DOES THE GUIDELINE DOCUMENT MAKE MAJOR RECOMMENDATIONS?

☐ Yes (List below, and then go to Question 30)

☐ No (Go to Question 31)

ASSESSOR: Briefly list in the space below the recommendations from the guideline document that the developers consider major. If the developers have not specifically indicated which are their major recommendations, please list those that you have used in answering the questions about the strength of scientific evidence.

30. THE GUIDELINE DOCUMENT EXPLICITLY DISCUSSES THE STRENGTH OF THE SCIENTIFIC EVIDENCE ON WHICH *EACH MAJOR RECOMMENDATION* IS BASED.

☐ Yes (Go to Question 30.1)

☐ No (Go to Question 30.2)

30.1. THE DISCUSSION OF THE STRENGTH OF THE EVIDENCE ON WHICH EACH MAJOR RECOMMENDATION IS BASED IS:

☐ Satisfactory for all recommendations

☐ Conditionally satisfactory-i.e., satisfactory for some but not all recommendations

☐ Unsatisfactory-i.e., not satisfactory for most or all recommendations

Comments:

>> GO TO QUESTION 31 >>

30.2. OMISSION OF A DISCUSSION OF THE STRENGTH OF THE SCIENTIFIC EVIDENCE FOR *EACH MAJOR RECOMMENDATION* IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

31. EACH MAJOR RECOMMENDATION IS CONSISTENT WITH THE ESTIMATED BENEFITS, HARMS, AND COSTS OF THE SERVICE OR INTERVENTION (AND THUS WITH THE STRENGTH OF THE EVIDENCE).

☐ Yes, completely ☐ Yes, partially ☐ No

Comments:

Potential Conflict Among Similar Sets of Guidelines

32. THE GUIDELINE DOCUMENT IDENTIFIES OTHER SETS OF GUIDELINES THAT DEAL WITH THE SAME CLINICAL CONDITION, TECHNOLOGY, OR TOPIC.

☐ Yes (Go to Question 33)

☐ No, but similar sets of guidelines are known to exist (Specify below and go to Question 33.2)

☐ Not applicable, no similar sets of guidelines are known to exist (Go to V. CLARITY)

33. THE GUIDELINE DOCUMENT IDENTIFIES POSSIBLE CONFLICTS AMONG EXISTING GUIDELINES AND THE REASONS FOR THEM.

☐ Yes (Go to Question 33.1) ☐ No (Go to Question 33.2)

33.1. THE DISCUSSION OF POSSIBLE CONFLICTS AMONG GUIDELINES IS:

☐ Satisfactory Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

>> GO TO V. CLARITY >>

33.2. OMISSION OF A DISCUSSION OF SIMILAR GUIDELINES, OR OF POSSIBLE CONFLICTS AMONG GUIDELINES, IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

V. CLARITY

Clarity means that guidelines are written in unambiguous language and terms, that the logic of the recommendations is clear and straightforward, and that the guideline document has a clear and easy-to-understand structure and format. That is, clarity encompasses the language and the logic with which the guideline document is written and the way it is physically presented. Clarity applies to three content areas of guidelines: (1) a general framework in which health condition(s), health practice(s), patient care goals, and similar topics are defined and discussed; (2) presentation and discussion of the evidence used in developing the guidelines; and (3) recommendations.

More specifically, this attribute requires that, as described below, certain things about language and terms, logic, and structure must be true.

Language and Terms

The guidelines are written in unambiguous language. Vague terms are avoided when describing the patient populations, health conditions, the health interventions, and the recommendations. For example, expressions such as "severe bleeding" are avoided in favor of (or at least qualified by) more precise language, such as a "drop in hematocrit of more than 6 percent in less than 8 hours." Or, for instance, a recommendation such as "thyroid function tests should be obtained whenever appropriate" is replaced by a recommendation that includes the type of test, its frequency, and the specific circumstances under which it should be used, such as "once every 5 years in otherwise healthy adults more than 65 years of age."

34. THE GUIDELINES DESCRIBE THE HEALTH CONDITION TO BE PREVENTED, DETECTED, OR TREATED IN UNAMBIGUOUS TERMS.

☐ Yes ☐ No

Comments:

35. THE GUIDELINES DESCRIBE THE OPTIONS FOR MANAGEMENT OF THE HEALTH CONDITION (I.E., THE HEALTH PRACTICE AND ITS ALTERNATIVES) IN UNAMBIGUOUS TERMS.

☐ Yes ☐ No

Comments:

36. IF THE GUIDELINES GIVE MAJOR RECOMMENDATIONS, EACH IS WRITTEN IN UNAMBIGUOUS TERMS.

ASSESSOR: Refer to the list you developed for Question 29 in answering this question.

☐ Yes ☐ No ☐ Not applicable, no major recommendation given

Comments:

Logic

The guidelines are as comprehensive as possible in keeping with the attributes "clinical adaptability" and "clinical flexibility." Thus, the logic of the guidelines is such that all clinically important and relevant situations are handled in a consistent, reasonable, and easy-to-follow manner and that situations that are not covered are explained in a logically appropriate place in the guideline statement.

Recommendations are mutually exclusive; that is, they are consistent with each other. For example, a guideline does not recommend "aortic valvuloplasty, for an 80-year-old man with end stage renal disease" in one place and "aortic valve replacement for an 80-year-old man with end stage renal disease" in another.

37. RECOMMENDATIONS ARE COMPREHENSIVE, INsofar AS THE EVIDENCE PERMITS, AND RECOMMENDATIONS THAT MIGHT BE EXPECTED ARE

GIVEN. (That is, the recommendations collectively cover all clinically relevant circumstances.)

☐ Yes (Go to Question 38) ☐ No (Go to Question 37.1)

Comments:

37.1. IF EXPECTED RECOMMENDATIONS SEEM TO BE MISSING, THE GUIDELINE DOCUMENT DISCUSSES WHY.

☐ Yes ☐ No

Comments:

38. RECOMMENDATIONS ARE CONSISTENT. (That is, no two recommendations in the guidelines conflict with each other.)

☐ Yes

☐ No (at least two recommendations appear to conflict with each other)

☐ Not applicable, no recommendations given

Comments:

Structural Clarity

The overall organization and appearance of the guideline document and the mode of presentation of the recommendations are easy for users to understand and follow. A structurally clear guideline is one in which the recommendations are easily accessible to

the prospective user. That is, clinicians should not have to read, analyze critically, and distill a detailed manuscript in order to find needed recommendations. Structural clarity may be achieved through the use of a summary, special highlighting techniques, algorithms, or other methods.

39. THE GUIDELINE DOCUMENT USES CLEAR HEADINGS, INDEXES, LISTS, FLOW CHARTS, OR OTHER DEVICES TO IDENTIFY MAJOR TOPICS DISCUSSED.

Comments:

40. THE GUIDELINE DOCUMENT HAS A SUMMARY OR ABSTRACT THAT ACCURATELY REFLECTS THE METHODS, CONTENT, AND RECOMMENDATIONS OF THE ENTIRE DOCUMENT.

☐ Yes ☐ No

Comments:

41. A USER OF THE GUIDELINE DOCUMENT CAN EASILY FIND EACH MAJOR RECOMMENDATION.

ASSESSOR: Refer to the list developed for Question 29 in answering this question.

☐ Yes ☐ No ☐ Not applicable, no major recommendation given

Comments:

VI. SCHEDULED REVIEW

Scheduled review means that a statement specifying a date for review and possible revision of the guideline has been included in the guideline document. Revisions to guidelines should reflect new clinical evidence or changing professional consensus.

This attribute requires that the guideline document either (1) give a specific date for review and possible revision of the guidelines or (2) describe a process by which such a date might be established and the review and possible revision performed.

42. THE GUIDELINE DOCUMENT GIVES A SPECIFIC DATE FOR SCHEDULED REVIEW, GIVES OTHER INFORMATION CONCERNING A PROCEDURE BY WHICH SCHEDULED REVIEW MIGHT BE DONE, OR GIVES A SUNSET OR EXPIRATION DATE.

☐ Yes (Go to Question 42.1) ☐ No (Go to Question 42.2)

42.1. ASSESSOR: Respond to Items 42a-42d, below, to determine whether the scheduled review date information is satisfactory, then answer Question 42.1 below, using your best judgment as to the overall rating for this attribute of scheduled review. Other factors you judge important should be specifically recorded under "Comments or Other Factors."

42a. The target date for review is:

☐ Appropriate ☐ Inappropriate ☐ None given/discussed

42b. The rationale for the target date is:

☐ Appropriate ☐ Inappropriate ☐ None given/discussed

42c. The procedures suggested for determining when the guidelines should be reviewed are:

☐ Appropriate ☐ Inappropriate ☐ None given/discussed

42d. The guideline has a sunset provision that may dictate when a scheduled review should take place or that may indicate when the guideline will expire.

☐ Appropriate ☐ Inappropriate ☐ None given/discussed

Now answer:

42.1. THE SCHEDULED REVIEW DATE OR PROCEDURE FOR SETTING IT IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments or Other Factors:

>> Go to VII. MULTI-DISCIPLINARY PROCESS >>

42.2. THE LACK OF A SCHEDULED REVIEW DATE OR PROCEDURE FOR SETTING IT IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

VII. MULTI-DISCIPLINARY PROCESS

A *multi-disciplinary process* for practice guidelines means that representatives of a broad range of practitioners, consumers or patients, and other groups likely to be affected by the guidelines have participated in the development process at some stage. These representatives can be individuals who have had direct responsibility for the guideline document or individuals who have reviewed that document or in other ways have contributed to it. This attribute intends that both methodologic and clinical disciplines be involved in the guideline-development process. This document cannot identify in advance all relevant participants, interested parties, or disciplines because each set of guidelines will differ in this respect.

This attribute requires that five things be true. First, some combination of individuals directly responsible for guidelines and those who have otherwise contributed to their development collectively represents all the key groups likely to affect or to be affected by the guidelines. Second, the guideline document describes the parties involved (including their credentials and potential biases); "the parties involved" is understood to mean participants in the actual development panel and those in review panels, public hearings, or other review forums. Third, potential biases and conflicts of interests have been discussed or otherwise appropriately taken account of. Fourth, the methods used to solicit panelists' views and arrive at group judgments have been described and are adequate and appropriate to the task of balancing views and potential biases. Fifth, the methods used to solicit outside review comments and present those to panelists have been described and are adequate to the task of making outside views clear to panelists.

43. PERSONS WITH APPROPRIATE CLINICAL AND METHODOLOGIC DISCIPLINES PARTICIPATED IN DEVELOPING THE GUIDELINE DOCUMENT--- THAT IS, A MULTI-DISCIPLINARY APPROACH WAS FOLLOWED.

___ Yes (Go to Question 43.1)

___ No (Go to Question 43.2)

___ Don't know or can't tell (Go to Question 43.2)

43.1. ASSESSOR: Respond to Items 43a-43i, below, to determine whether the multi-disciplinary process is satisfactory; then answer Question 43.1 below, using your best judgment as to the overall rating for this element of multi-disciplinary process. Other factors you judge important should be specifically recorded under "Comments or Other Factors."

43a. An explanation, discussion, or rationale for selecting the *guideline panel chairperson* is given.

☐ Yes ☐ No

43b. An explanation, discussion, or rationale for selecting the members of the guideline panel is given.

☐ Yes ☐ No

43c. An explanation, discussion, or rationale for selecting other individuals directly responsible for the guideline document (such as consultants) is given.

☐ Yes ☐ No

43d. The explanation(s), discussion(s), or rationale(s) for selecting the individuals covered in 43a-c is (are):

☐ Adequate ☐ Inadequate ☐ Not applicable

43e. These individuals reflect all appropriate interest groups and disciplines.

☐ Yes ☐ No ☐ Can't tell

43f. One or more outside review panel(s) commented on or reviewed draft guidelines.

☐ Yes ☐ No ☐ Can't tell

43g. One or more public hearing(s) or similar review mechanism(s) were held to allow comment or review on draft guidelines.

☐ Yes ☐ No ☐ Can't tell

43h. Collectively, the review panel(s), public hearing(s), or other review mechanisms reflected all appropriate interest groups and disciplines.

☐ Yes ☐ No ☐ Can't tell

43i. If the answer to either question 43e or question 43h is "No" or "Can't tell," please record what groups or disciplines appear to have been omitted.

Now answer:

43.1. THE MULTI-DISCIPLINARY APPROACH TO THE GUIDELINES DEVELOPMENT PROCESS IS:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments or Other Factors:

43.2. THE LACK (OR APPARENT LACK) OF A MULTIDISCIPLINARY PROCESS IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

44. THE GUIDELINE DOCUMENT EXPLICITLY NOTES ANY POTENTIAL BIASES AND/OR CONFLICTS OF INTERESTS OF THE PANEL MEMBERS, OR STATES THAT BIASES AND CONFLICTS OF INTEREST WERE DISCUSSED AMONG PANEL MEMBERS OR OTHERWISE TAKEN INTO ACCOUNT.

☐ Yes, potential biases and/or conflicts of interest are noted

☐ Yes, a statement that biases and/or conflicts of interest were discussed is given

☐ No, no note or statement about biases and/or conflicts of interest is given

45. OVERALL, POTENTIAL BIASES AND/OR CONFLICTS OF INTEREST APPEAR TO BE ADEQUATELY BALANCED OR OTHERWISE ACCOUNTED FOR IN THE GUIDELINE DEVELOPMENT PROCESS.

☐ Yes ☐ No (Specify) ☐ Don't know or can't tell

Comments:

46. THE GUIDELINE DOCUMENT DESCRIBES THE METHODS USED TO SOLICIT VIEWS OF INTERESTED PARTIES NOT ON THE GUIDELINES DEVELOPMENT PANEL AND TO PRESENT THOSE VIEWS TO THE MEMBERS OF PANEL.

☐ Yes (Go to Question 46.1) ☐ No (Go to Question 46.2)

46.1. THE METHODS USED TO SOLICIT VIEWS OF THOSE NOT ON THE PANELS AND PRESENT THOSE VIEWS TO PANELS ARE:

☐ Satisfactory ☐ Conditionally satisfactory ☐ Unsatisfactory (Specify)

Comments:

46.2. THE LACK OF A DESCRIPTION OF THE METHODS USED TO SOLICIT VIEWS OF THOSE NOT ON THE PANELS AND TO PRESENT THOSE VIEWS TO PANELS IS:

☐ Unimportant ☐ Minor omission ☐ Major omission

Comments:

PLEASE RECORD ANY SUMMARY JUDGMENTS OR OTHER COMMENTS YOU MAY HAVE AND ANY RECOMMENDATIONS FOR ADDITIONAL REVIEW.

PART THREE. SUMMARY EVALUATION SHEET

Instructions and Key

ASSESSOR: Upon completing the entire assessment instrument, please record answers to the main questions (Questions 1-46) below. Circle the relevant answer, according to the following key:

KEY

Y = Yes; YQ = yes, but response qualified;

N = No; NQ = no, but response qualified;

S = Satisfactory, CS = Conditionally satisfactory, US = Unsatisfactory;

UN = Unimportant, MI = Minor omission, MA = Major omission;

NA = Not applicable

DK = Don't know, or can't tell

I. CLINICAL APPLICABILITY

Y	N		1. Description of patient population
S	CS	US	1.1. Quality of description
UN	MI	MA	1.2. Omission of description
Y	N	NA	2. Discussion of complex clinical problems
S	CS	US	2.1. Quality of discussion
UN	MI	MA	2.2. Omission of discussion
Y	N		3. Rationale for excluding patient populations
S	CS	US	3.1. Quality of rationale
UN	MI	MA	3.2. Omission of rationale

II. CLINICAL FLEXIBILITY

Y	N		4. Information about acceptable clinical exceptions
S	CS	US	4.1. Quality of information or statement
UN	MI	MA	4.2. Omission of information or statement
Y	N		5. Information about acceptable nonclinical exceptions
S	CS	US	5.1. Quality of information or statement
UN	MI	MA	5.2. Omission of information or statement
Y	N		6. Discussion of patient preferences in the health care decisions

S	CS	US	6.1. Quality of discussion
UN	MI	MA	6.2. Omission of discussion
Y	N		7. Discussion of patient preferences in guideline development
S	CS	US	7.1. Quality of discussion
UN	MI	MA	7.2. Omission of discussion

III. RELIABILITY/REPRODUCIBILITY

Y	N		8. Independent review by experts or outside panels
S	CS	US	8.1. Quality of discussion
Y	N		9. Explanation of lack of independent review
S	CS	US	9.1. Quality of explanation
UN	MI	MA	9.2. Omission of explanation
Y	N		10. Guidelines pretested in some manner
S	CS	US	10.1. Quality of discussion
Y	N		11. Explanation of lack of pretesting
S	CS	US	11.1. Quality of explanation
UN	MI	MA	11.2. Omission of explanation

IV. VALIDITY

STRENGTH OF THE SCIENTIFIC EVIDENCE AND PROFESSIONAL CONSENSUS

Y	N		12. Method of collecting (identifying and retrieving) scientific evidence is specifically described
S	CS	US	12.1. Quality of method
UN	MI	MA	12.2. Lack of method
Y	N		13. Adequate references to sources of scientific evidence
Y	N		14. General discussion of strength of scientific evidence
Y	N		15. Explicit rating of the strength of the scientific evidence
S	CS	US	15.1. Quality of rating method
UN	MI	MA	15.2. Lack of general discussion of rating method

Y	N	NQ	16. If a formal method of synthesis is used, explicit description of the method
S	CS	US	16.1. Quality of formal method
UN	MI	MA	16.2. Omission of description of formal method
Y	N		17. If applicable, the results of a formal synthesis of scientific evidence are explicitly reported
S	CS	US	17.1. Quality of results of the synthesis
UN	MI	MA	17.2. Omission of results of the synthesis
Y	N		18. If applicable, the expert or group judgment techniques used for reaching professional consensus are explicitly described
S	CS	US	18.1. Quality of expert or group judgment techniques
UN	MI	MA	18.2. Omission of description of expert or group judgment techniques
Y	N		19. If applicable, the strength of professional consensus resulting from use of group judgment techniques is reported
S	CS	US	19.1. Quality of information about strength of professional consensus
UN	MI	MA	19.2. Omission of explicit information about strength of professional consensus

HEALTH BENEFITS AND HARMS/RISKS: QUALITATIVE DESCRIPTION

Y	N		20. Qualitative description of health benefits
U	S	US	20.1. Quality of qualitative description
UN	MI	MA	20.2. Omission of qualitative description
Y	N		21. Qualitative description of potential harms or risks
S	CS	US	21.1. Quality of qualitative description
UN	MI	MA	21.2. Omission of qualitative description

HEALTH BENEFITS AND HARMS/RISKS: QUANTITATIVE INFORMATION

Y	N		22. Quantitative information or estimates of health benefits
S	CS	US	22.1. Quality of quantitative information
UN	MI	MA	22.2. Omission of quantitative information

Y	N	NA	23. Health benefits projected in terms of life expectancy or similar measures
Y	N		24. Quantitative information or estimates of potential harms or risks
S	CS	US	24.1. Quality of quantitative information
UN	MI	MA	24.2. Omission of quantitative information

HEALTH COSTS: QUALITATIVE DESCRIPTION

Y	N		25. Qualitative description of health costs or expenditures
S	CS	US	25.1 Quality of qualitative description
UN	MI	MA	25.2. Omission of qualitative description

HEALTH COSTS: QUANTITATIVE DESCRIPTION

Y	N		26. Quantitative information or estimates of health costs or expenditures
S	CS	US	26.1. Quality of quantitative information
UN	MI	MA	26.2. Omission of quantitative information
Y	N	NA	27.If health benefits projected in terms of life expectancy or similar measures, costs per unit of each identified benefit also estimated
Y	YQ	N	28. Generally, estimates of benefits, harms, and costs are consistent with the strength of provided evidence
Y	N		29. Major recommendations made in the guideline
Y	N		30. Discussion of strength of the scientific evidence for each major recommendation
S	CS	US	30.1. Quality of discussion
UN	MI	MA	30.2. Omission of discussion
Y	YQ	N	31. Each major recommendation consistent with strength of scientific evidence

POTENTIAL CONFLICT AMONG SIMILAR SETS OF GUIDELINES

Y	N	NA	32. Other sets of guidelines identified
Y	N		33. Possible conflicts among existing guidelines discussed
C	CS	US	33.1. Quality of discussion
UN	MI	MA	33.2. Omission of discussion

V. CLARITY

LANGUAGE AND TERMS

Y	N		34. Language describing the health condition is unambiguous
Y	N		35. Language describing the options for management is unambiguous
Y	N	NA	36. Language for each major recommendation is unambiguous

LOGIC

Y	N		37. Recommendations are comprehensive and present when expected
Y	N		37.1. Reasons given for lack of expected recommendations
Y	N	NA	38. Recommendations are consistent

STRUCTURAL CLARITY

Y	N		39. Guideline document uses clear headings, indexes, etc.
Y	N		40. Guideline document has accurate summary or abstract
Y	N	NA	41. Users can find recommendations easily

VI. SCHEDULED REVIEW: DEFINITION AND EVALUATION QUESTIONS

Y	N		42. Scheduled date for review or a procedure for arriving at such a date is provided
S	CS	US	42.1. Quality of the scheduled review date or procedure for setting one
UN	MI	MA	42.2. Lack of a scheduled review date or procedure for setting one

VII. MULTI-DISCIPLINARY PROCESS

Y	N	DK	43. Participation of persons in appropriate clinical and methodologic disciplines
---	---	----	---

S	CS	US	43.1. Quality of the multi-disciplinary approach
UN	MI	MA	43.2. Lack of a multi-disciplinary process
Y	YQ	N	44. Guideline document notes potential biases or conflicts of interest or indicates they were taken into account
Y	N	DK	45. Balance of potential biases or conflicts of interest
Y	N		46. Description of the methods used to solicit views of those not on the guidelines development panel and to present those views to the panel
S	CS	US	46.1. Quality of methods used
UN	MI	MA	46.2. Lack of a description of methods used

SUMMARY JUDGMENT, OTHER COMMENTS, OR NEED FOR ADDITIONAL REVIEW:

This page intentionally left blank.

APPENDIX C: CDC/ATSDR FEDERALLY CHARTERED ADVISORY COMMITTEES *

Office of the Director

- ▶ Advisory Committee on Immunization Practices (ACIP)
 1. Dixie E. Snider, M.D., M.P.H.
 2. Gloria Kovach, Committee Management Specialist (639-3851/Fax 639-3036)
- ▶ Advisory Committee to the Director, CDC (ACD, CDC)
 1. Linda Kay McGowan (Acting)
 2. Marge Carter, Program Specialist (639-2290/Fax 639-3941)
- ▶ CDC Advisory Committee on the Prevention of HIV Infection (CDC, ACPHI)
 1. Helene D. Gayle, M.D. (Acting)
 2. Connie Granoff, Program Specialist (639-8029/Fax 8600)
- ▶ National Vaccine Advisory Committee (NVAC)
 1. Robert F. Breiman, M.D.
 2. Gloria Kovach, Committee Management Specialist (639-3851/Fax 639-3036)

National Center for Chronic Disease Prevention and Health Promotion

- ▶ Breast and Cervical Cancer Early Detection and Control Advisory Committee (BCCEDCAC)
 1. Rebecca B. Wolf
 2. Karen Norton (488-4751/Fax 488-4727)
- ▶ Interagency Committee on Smoking and Health (ICSH)
 1. Karen M. Deasy
 2. Amy Garson, Staff Specialist (202-205-8500/Fax 202-205-8313)
- ▶ Technical Advisory Committee for Diabetes Translation and Community Control Programs (TACDTCCP)
 1. Frank Vinicor, M.D., M.P.H.
 2. Cheryl Shaw, Program Specialist (488-5004/Fax 488-5966)

National Center for Environmental Health

- ▶ Advisory Committee for Energy-Related Epidemiologic Research (ACERER)
 1. Richard J. Jackson, MD, MPH
 2. Nadine Dickerson, Program Specialist (488-7040/Fax 488-7044)

- ▶ Advisory Committee on Childhood Lead Poisoning Prevention (ACCLPP)
 1. Henry Falk, M.D.
 2. Barbara Nelson, Program Analyst (488-7330/Fax 488-7335)
- ▶ Hanford Thyroid Morbidity Study Advisory Committee (HTMSAC)
 1. Henry Falk, M.D.
 2. Nadine Dickerson, Program Specialist (488-7040/Fax 488-7044)

National Center for Injury Prevention and Control

- ▶ Advisory Committee for Injury Prevention and Control (ACIPC)
 1. Thomas E. Blakeney (Acting)
 2. Iris Lansing, Program Specialist (488-4821/Fax 488-4338)
- ▶ Injury Research Grant Review Committee (IRGRC)
 1. Richard W. Sattin, M.D., F.A.C.P.
 2. Iris Lansing, Committee Management Specialist (488-4821/Fax 488-4338)

National Center for Health Statistics

- ▶ National Committee on Vital and Health Statistics (NCVHS)
 1. Gail F. Fisher, Ph.D.
 2. Jackie Adler, Conference Assistant (301-436-7122/Fax 301-436-4233)

National Center for Infectious Diseases

- ▶ Board of Scientific Counselors, National Center for Infectious Diseases (BSC, NCID)
 1. Rosemary B. Ramsey
 2. Diane Holley, Committee Management Specialist (639-0078/Fax 639-3853)
- ▶ Hospital Infection Control Practices Advisory Committee (HICPAC)
 1. Julia S. Garner, R.N., M.N.
 2. Karen Friend, Secretary (HIP) (639-6403/Fax 639-6458)

National Center for Prevention Services

- ▶ Advisory Council for the Elimination of Tuberculosis (ACET)
 1. Helene Gayle, M.D.
 2. Tracy Whitnell, Program Analyst (639-8006/Fax 639-8600)

National Institute for Occupational Safety and Health

- ▶ Board of Scientific Counselors, National Institute for Occupational Safety and Health
(BSC, NIOSH)
 1. Bryan D. Hardin, Ph.D. (Acting)
 2. Judy James, Committee Management Specialist (639-4403/Fax 639-2196)
- ▶ Mine Health Research Advisory Committee (MHRAC)
 1. Gregory R. Wagner, M.D.
 2. Judy James, Committee Management Specialist (639-3794/Fax 639-2196)
- ▶ Safety and Occupational Health Study Section (SOHSS)
 1. Contact: Roy M. Fleming, Sc.D.
 2. Judy James, Committee Management Specialist (639-3794/Fax 639-2196)
- ▶ Workers' Family Protection Task Force (WFPTF)
 1. Elizabeth A. Whelan, Ph.D.
 2. Pam Graydon (513/533-8312)

Public Health Practice Program Office

- ▶ Clinical Laboratory Improvement Advisory Committee (CLIAC)
 1. Edward L. Baker, M.D.
 2. Julie Wasil, Committee Management Specialist (488-4651/Fax 488-7667)

ATSDR Federal Advisory Committee

- ▶ Board of Scientific Counselors, Agency for Toxic Substances and Disease Registry
(BSC, ATSDR)
 1. Charles Xintaras, Sc.D.
 2. Diane Allgood, Secretary (639-0708/Fax 639-0586)

Federal Advisory Committees Utilized by CDC and ATSDR

- ▶ Citizens Advisory Committee on Public Health Service Activities and Research at Department of Energy Sites (CACPHSARDES)
 - Fernald
 1. Steve Adams
 2. Nadine Dickerson, Program Specialist (488-7040) (Fernald)
 - Hanford Health Effects Subcommittee (HHES)

1. Jim Carpenter
 2. Linda Carnes (639-0730/Fax 639-0759) (Hanford Health Effects Subcomm - HHES)
- Idaho National Engineering Laboratory (INEL)
 1. Arthur J. Robinson, Jr.
 2. Nadine Dickerson, Program Specialist (488-7040) (Idaho National Engineering Lab - INEL)
 - Savannah River Site (SRS)
 1. Paul Renard
 2. Nadine Dickerson, Program Specialist (488-7040/Fax 488-7044) (Savannah River Site - SRS)
- Disease, Disability, and Injury Prevention and Control Special Emphasis Panel
(Panel name determined by program conducting the SEP)
1. Burma Burch (CDC and ATSDR Contact)
 2. Judi Cook, Committee Management Specialist (639-6389/Fax 639-6290)
- * **Most of these committees develop guidelines: 1 = executive secretary and 2 = management contact**

APPENDIX D: ACCESSING CDC PREVENTION GUIDELINES DATABASE

The Centers for Disease Control and Prevention maintains an electronic data base of recommendations and guidelines for disease and injury prevention in both clinical care and public health practice. With over 400 documents, the CDC Prevention Guidelines Data Base (PGDB) includes guidelines for the prevention and control of a wide range of conditions relevant to public health and clinical prevention. Some of the guideline topics are AIDS, cholera, disaster response, dengue fever, suicide, lung cancer, vaccine-preventable diseases, sexually transmitted diseases, birth defects, and malaria. About two-thirds of the documents in the data base are in their original form as published in CDC's Morbidity and Mortality Weekly Report (MMWR). The others were published as CDC monographs, books or book chapters, brochures, or articles in peer-reviewed journals. Some of these other publications are not contained within the data base; however, a contact is identified so that the document can be obtained from CDC. The PGDB does include all of the ACIP immunization recommendations, the Guide to Clinical Preventive Services, CDC treatment guidelines for sexually transmitted diseases, and the entire "yellow book" (Health Information for International Travelers).

The PGDB is now widely accessible through the World Wide Web via CDC's home page (<http://www.cdc.gov/cdc.html>) or directly (<http://wonder.cdc.gov/wonder/prevguid/prevguid.html>). The PGDB is also available on CD-ROM as a stand-alone version with its own search engine. The main data base will be updated weekly and the CD-ROM version will be updated quarterly. The PGDB is accessible within CDC as indicated above. Additional information concerning the PGDB can be obtained through CDC WONDER Customer Support at (888)496-8347; or E-mail cwus@cdc.gov.