Consensus Development at the NIH: Improving the Program

Report of a Study by a Committee of the Institute of Medicine, Council on Health Care Technology

ISBN: 0-309-58258-X, 96 pages, 6 x 9, (1990)

Consensus Development at the NIH: Improving the Program

Report of a Study by a Committee of the Institute of Medicine
Council on Health Care Technology

NATIONAL ACADEMY PRESS
WASHINGTON, D.C. 1990

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The Institute of Medicine was chartered in 1970 by the National Academy of Sciences to enlist distinguished members of the appropriate professions in the examination of policy matters pertaining to the health of the public. In this, the Institute acts under both the Academy's 1863 congressional charter responsibility to be an adviser to the federal government and its own initiative in identifying issues of medical care, research, and education.

Support for this project was provided by the National Institutes of Health, Office of Medical Applications of Research, pursuant to Contract No. NO1-OD-8-2110.

Library of Congress Catalog Card Number 90-60935
International Standard Book Number 0-309-04242-9
Publication IOM-90-04

Additional copies of this report are available from:

National Academy Press
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

S131

Printed in the United States of America

Acknowledgments

The committee thanks the individuals who provided insight into the National Institutes of Health Consensus Development Program. In particular, the committee acknowledges Donald S. Fredrickson, former director of the National Institutes of Health; John Ferguson, director of the Office of Medical Applications of Research; Jerry Elliott, program analyst, Office of Medical Applications of Research; and Jacqueline Kosecoff, executive vice president of Value Heath Sciences, Inc., and professor of medicine and public health, University of California at Los Angeles.

The committee would like to express its appreciation of the Institute of Medicine staff who facilitated the work of this committee, especially Elaine Steen, Clifford Goodman, Sharon Baratz, and Ethan Halm. We are grateful to Holly Dawkins and Evanson Joseph for logistical support and help in preparing this manuscript.

The committee also thanks the members of the Council on Health Care Technology and the Institute of Medicine Council, who reviewed and provided useful comments on the report.

COMMITTEE TO IMPROVE THE NATIONAL INSTITUTES OF HEALTH CONSENSUS DEVELOPMENT PROGRAM

COUNCIL ON HEALTH CARE TECHNOLOGY

Chair

J. SANFORD SCHWARTZ
Executive Director, Leonard Davis Institute of Health Economics, Associate Professor of Medicine and Health Care Systems, and Robert D. Eilers Associate Professor of Health Care Management and Economics, University of Pennsylvania

MARCIA ANGELL
Executive Editor, ***The New England Journal of Medicine***

ROBERT H. BROOK*
Deputy Director, Health Sciences Program, The RAND Corporation, and Chief, Division of Geriatrics and Professor of Medicine and Public Health, University of California at Los Angeles

FLOYD W. DENNY*
Professor of Pediatrics, School of Medicine, University of North Carolina at Chapel Hill

ROBERT M. DONALDSON, JR.
Deputy Dean, Yale University School of Medicine

DENNIS **FRYBACK**
Professor of Industrial Engineer; ing and of Preventive Medicine, University of Wisconsin-Madison

ARNOLD D. KALUZNY
Professor of Health Policy and Administration, School of Public Health, University of North Carolina at Chapel Hill

R. BRIAN HAYNES
Professor of Clinical Epidemiology and Medicine, **McMaster** University

JONATHAN LEFF
Editorial Consultant, Former Director, Consumer Reports Books, South Hero, Vermont

STEPHEN G. PAUKER
Professor and Chief, Division of Clinical Decision Making, Department of Medicine, New England Medical Center, Tufts University School of Medicine

CAROL K. REDMOND
Professor and Chair, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

* Member, Institute of Medicine

INSTITUTE OF MEDICINE

SAMUEL O. THIER, President
ENRIQUETA C. BOND, Executive Officer

Project Staff

CLIFFORD GOODMAN, Director, Council on Health Care Technology
ELAINE B. STEEN, Project Officer
SHARON R. BARATZ, Research Associate
HOLLY DAWKINS, Project Assistant
EVANSON JOSEPH, Project Assistant

Preface

One of the goals of the Institute of Medicine's Council on Health Care Technology has been to promote the development and application of technology assessment in health care. Assessment of medical technologies currently relies extensively on group judgment or "consensus development" processes. These methods are used by medical professional societies, private and public third-party payers, biomedical research agencies, and others to assess state-of-the-art medical and surgical procedures and other technologies, to define standard and accepted medical practices, to bridge gaps and resolve disparities among research findings, and to establish coverage and reimbursement policies. Despite the often weighty implications for health policy of decisions based on group judgment, these processes often are applied with little systematic or critical examination of their conduct, utility, or appropriateness for technology assessment.

The panels convened for group judgment must weigh, integrate, and interpret the available evidence, experience, beliefs, and values in order to formulate guidelines, recommendations, or other findings. The technologies assessed may be evolving rapidly; the evidence available to the panels may consist of a sparse patchwork of research results of varying quality. Panelists may be subject to biases and errors of reasoning; experts and nonexperts alike may be

subject to oversimplification, empiricism, case-selection biases, incentives, and advocacy.

Group judgment methods have been developed and used in many fields, including the social, political, physical, and biomedical sciences; engineering; defense; environmental studies; and other domains. To improve group judgment for assessing medical technologies, the council urges programs to learn from alternative techniques and models used in other fields for engaging groups of experts in rendering well-founded and informed findings.

The council is conducting a three-part examination of group judgment methods for assessing medical technologies: an international workshop on consensus development for medical technology, a study to improve the National Institutes of Health (NIH) Consensus Development Program, and a workshop to improve group judgment for medical practice and technology assessment. First, in conjunction with the June 1989 annual meeting of the International Society for Technology Assessment in Health Care, the council coordinated a workshop that compared and contrasted national consensus development programs. The proceedings of the workshop (*Improving Consensus Development for Health Technology Assessment: An International Perspective*, C. Goodman and S. Baratz, eds., 1990), in which representatives of 11 countries participated, include a consolidated set of participants' recommendations for improving these approaches. Second, as requested by the NIH Office of Medical Applications of Research, a committee operating under the aegis of the council conducted a November 1989 meeting and follow-up work to review and provide recommendations for improving the NIH Consensus Development Program. The committee's recommendations are presented in this report. Third, a May 1990 workshop sponsored by the council will convene those who conduct, participate in, and use the results of group judgment efforts of medical professional societies, third-party payers, regulatory agencies, and others. This workshop will focus on improving key methodologic aspects of these programs and sharing newer approaches to improve their application for rendering practice guidelines, coverage decisions, and other assessment-related findings and policies.

The council's goal in undertaking this three-part effort is to improve group judgment for assessing medical technologies. This report makes a significant contribution toward that goal. Although written specifically for the NIH Consensus Development Program,

the recommendations presented in this report address many of the challenges of any group process intended to produce a consolidated, well-substantiated expert judgment for direct application to policymaking. Furthermore, by requesting this evaluation, NIH may encourage others in the field to undertake efforts to examine and improve their group judgment efforts.

Contents

Summary

At the request of the National Institutes of Health (NIH), the Institute of Medicine established the Committee on Improving the National Institutes of Health Consensus Development Program (CDP). The committee's report provides specific recommendations to the Office of Medical Applications of Research (OMAR) for strengthening the Consensus Development Program.

Four main themes arise from the committee's recommendations that address the structure and function of the NIH CDP. First, as the intent of the NIH CDP, as well as other technology assessment programs, is to change the behavior of clinicians and others in the health care system, the purpose, scope of inquiry, and planning of the program should reflect greater emphasis on understanding and representing the concerns of the users of the consensus statements. The scope of inquiry of the NIH CDP should be broadened to include relevant economic, social, and ethical aspects of assessing biomedical technologies. OMAR should expand the program's purpose to acknowledge explicitly that the ultimate goal of the program is to change behavior toward appropriate use of health practices and technologies. The program planning process should seek to identify the concerns of the broader health care system (i.e., beyond NIH) through solicitation of suggestions for conference topics, questions, panelists, and speakers.

Second, adding greater structure to certain elements of the NIH CDP will enhance the program's ability to generate useful findings.

The topic selection process should be formalized and should include consideration of evidence on the state of clinical practice. The planning committee should include members who understand the clinical, methodologic, and societal issues related to the conference topic and should reflect representation of the health care system's concerns about the topic. Conference questions should be reviewed by persons not involved in the process of drafting to ensure that the questions are clearly stated. Conference speakers should submit more thorough information in advance of the conference. When available data make it feasible, meta-analysis should be performed prior to convening the conference to obtain more rigorous interpretation of multiple sources of evidence. Evidence should be presented to the panel in a standardized format and be graded for its quality. A working definition of consensus should be made available to, and agreed upon by, the panel prior to initiating deliberations; the definition of consensus used by the panel should be included in the consensus statement. The consensus statement should include well-reasoned minority opinions when they exist, note when minority opinions do not exist, identify items that lack adequate evidence for a judgment to be made, and identify specific areas for further research.

Third, as is the case with group judgment programs in general, the NIH CDP can benefit from continued experimentation and evaluation. OMAR should develop an explicit ongoing research effort to determine ways to improve the CDP and to monitor the impact of the program. OMAR should experiment with quantitative decision modeling and other resources and means to aid the efforts of panels to achieve well-founded consensus.

Fourth, adequate financial support and a strong organizational commitment at the highest levels of NIH are required to enable OMAR to perform its challenging role in technology assessment and transfer on behalf of NIH. OMAR's reporting relationship to the director of NIH should be reevaluated with respect to the CDP. OMAR should be an equal partner with the bureaus, institutes, and divisions in setting the agenda of the CDP. An external advisory council should be established to assist OMAR in setting its agenda, including provision of oversight to the NIH CDP and guidance in consensus conference topic selection. OMAR should revise its budget expectations for consensus development conferences in light of the recommendations of this report, which likely would require additional resources.

1

Introduction

CONTEXT OF THIS REPORT

Since 1977 the National Institutes of Health (NIH) has conducted consensus development conferences to evaluate biomedical technologies and practices and has disseminated the results to health professionals and the public. The NIH Consensus Development Program (CDP), administered by the Office of Medical Applications of Research (OMAR) in NIH, is one of the more prominent health technology[1] assessment activities in the United States. The strengths of the NIH consensus development process

> are in its potential to translate a large body of research evidence into practical clinical policy, bring together apparently conflicting viewpoints, with the evidence as the "common denominator," draw public as well as professional attention to important clinical issues, obtain front-line practitioner input on the feasibility of evidence-generated clinical policy, and increase the exposure of all parties to the existing research evidence in an area" (Lomas, 1986).

The NIH CDP is widely known and has served as a model for the development of consensus development and group judgment programs in the United States and abroad (Andreasen, 1988; Calltorp,

[1] Throughout this report, *technology* includes drugs, devices, equipment, medical and surgical procedures, and other techniques and practices used in delivering health care and the systems in which such care is delivered.

1988; Goodman, 1988; Institute of Medicine [IOM], 1985; Lomas, 1986). Nevertheless, and perhaps because of the initiative of the program and its broad exposure, concerns about specific conferences and the program as a whole exist. Some observers have contended that data provided to certain panels have been incomplete, overinterpreted, or misrepresented (Perry, 1987); that a number of recommendations have been made in the absence of supporting evidence (Ahrens, 1985); that there is bias in the selection of panelists and speakers (Oliver, 1985); and that topics have not always been sufficiently important to warrant this type of evaluation or expenditure (Perry, 1987). The conference format, particularly the "grueling night sessions" (Mullan and Jacoby, 1985), has been described as onerous (Perry, 1987) and as perhaps leading to hurried conclusions (Oliver, 1985). Others have noted that minority views may have been obscured or submerged, that there has been too much focus on compromise between viewpoints given the available research evidence, that ambiguous or overly generalized recommendations have resulted because of poorly worded questions or excessive compromise, and that a vocal few have been able to dominate some proceedings (Lomas, 1986). Questions have been raised about the awareness of conclusions by clinicians (Jacoby, 1985) and the program's impact on physicians' behavior (Gleicher, 1984; Kanouse et al., 1989). Since the early 1980s OMAR has been actively involved in evaluating past conferences and incorporating the results of the analyses into its program guidelines. Although progress has been made, OMAR continues to seek improvements in the NIH CDP (IOM, 1985; Kanouse et al., 1989).

In response to a request from NIH to review the elements of its consensus development process and to develop recommendations to improve the process, IOM formed the Committee to Improve the National Institutes of Health Consensus Development Program under the aegis of the Council on Health Care Technology. The formal charge to the committee was to consider the following major components of the NIH process:

- topic selection
- role, size, and composition of the committee
- formulation of the agenda
- selection and preparation of speakers
- presentation and synthesis of conference data

- development of consensus
- dynamics of the committee
- preparation of initial and final consensus statements

The committee was also charged with formulating recommendations to improve the NIH CDP. The committee charge was specifically limited by the request of NIH not to address matters of the program's dissemination activity and impact, given already ongoing NIH activities in these areas. In particular, NIH contracted with Prospect Associates of Bethesda, Maryland, to organize an expert panel in January 1989 on the role of medical opinion leaders in disseminating consensus development conference recommendations. Also, the results of an evaluation of the NIH CDP's impact have been presented in a recent report by the RAND Corporation, *Changing Medical Practice Through Technology Assessment: An Evaluation of the National Institutes of Health Consensus Development Program* (Kanouse et al., 1989).

The committee found its charge to be narrow. Factors beyond program function (i.e., the components of the NIH CDP noted in the committee charge) affect the success of the program. Therefore, the committee's recommendations address program purpose and scope, organizational issues, and financial support in addition to program function. Further, dissemination of CDP findings is an essential element in the process, given the importance of using consensus findings to influence medical practice. Program impact should be examined when considering ways to improve the program. Thus, although the committee's circumscribed charge and the time available for the committee's work precluded it from specifically addressing the dissemination of consensus statements and impact of the process, the committee identified these as issues that require additional attention by NIH.

NIH CONSENSUS DEVELOPMENT PROGRAM

The NIH CDP[2] borrows from three models: (1) the judicial model, in which evidence is heard and weighed by knowledgeable impartial

[2] An extensive discussion of the NIH CDP is included in the background paper prepared for the committee that appears as Appendix A of this report.

judges or by juries of peers; (2) the scientific meeting, in which experts discuss their work with their peers; and (3) the town meeting, in which a forum is provided for all interested persons to express their views (Mullan and Jacoby, 1985). NIH consensus development conferences include participation by speakers who present evidence, an audience that has the opportunity to comment on the evidence, and a panel that deliberates and produces a written statement based on its judgment.

The planning and implementation process of a particular consensus development conference usually takes 12 to 15 months. It begins with selection of a topic from among suggestions typically originating from the bureaus, institutes, and divisions (BIDs) of NIH; other Public Health Service agencies such as the Food and Drug Administration or Centers for Disease Control; the Department of Veterans Affairs; the U.S. Congress; or organizations outside of government (Goodman, 1988). According to OMAR, four criteria are applied in evaluating potential topics: (1) the issue should have public health importance, (2) there should be controversy over scientific aspects of the issue, (3) there should be available evidence on which to base evaluation of the issue, and (4) the issue should be amenable to clarification on technical grounds (OMAR, NIH, 1988).

Once a conference topic is selected, a planning committee is formed consisting of OMAR staff, sponsoring institute staff, the prospective conference chair, and outside experts. The planning committee identifies key conference issues and drafts questions relating to the topic being assessed. These questions normally address issues of the safety and efficacy of the technologies at hand and define the dimensions of the conference. The planning committee also recommends conference panelists, program format, and speakers. Panel members are selected to represent the various areas of expertise necessary to address the conference questions (e.g., technologic, clinical, and methodologic expertise). Conference speakers present evidence about the technology being evaluated to the panelists, who are responsible for responding to the questions posed by the planning committee.

The conferences are open meetings to which members of the public and the medical community are invited. They usually last two and a half days. The first day and a half are normally spent in a plenary session in which speakers present information on the state of the science and the safety and efficacy of the technologies under consideration. These presentations are followed by open discussions engaging speakers, panelists, and members of the audience.

Following the plenary session, the panel convenes to consider the expert opinions of the conference speakers and other views expressed at the meeting and to draft responses to the conference questions. The resulting document, known as the consensus statement, is read to the audience on the morning of the third day for further comment and discussion among the panel and audience. The panel may choose to incorporate comments received during this session in the final consensus statement. The consensus statement may include minority opinions if a panel cannot achieve full agreement on a particular point. However, this option has been exercised only twice. The conference concludes with a press conference.

After a final consensus statement is approved by the panel, the document is published by OMAR and widely disseminated to health care providers and administrators, the biomedical research and education communities, and the general public. Conference reports and summaries are published in medical and other scientific journals pertinent to the conference topic. Most consensus statements are published in the *Journal of the American Medical Association,* and some are presented on dedicated medical television networks.

DESCRIPTION OF THIS STUDY

The IOM Committee to Improve the National Institutes of Health Consensus Development Program held a one-day meeting on November 21, 1989. It was preceded by an evening session on November 20 at which Donald S. Fredrickson, who served as director of NIH from 1975 to 1981, described the origin and development of the NIH CDP. The meeting agenda was organized to correspond with the following components of the NIH CDP:

- topic selection
- planning committee formation
- agenda formulation: question drafting and speaker selection
- consensus panel selection and composition
- preparation of speakers, data synthesis, and data presentation
- development of consensus and group dynamics
- initial and final consensus statement preparation

In addition to committee members, persons with experience in group judgment and with the NIH CDP were invited to participate in the meeting. Participants were provided with resource materials to help them prepare for the meeting.

The meeting included two presentations about the NIH CDP. John Ferguson, director of OMAR, provided an update on the NIH CDP. Jacqueline Kosecoff, executive vice president, Value Health Sciences, Inc., and adjunct professor of medicine and public health, University of California at Los Angeles, presented the relevant findings of the 1987 RAND Corporation study of the NIH CDP, *Changing Medical Practice Through Technology Assessment: An Evaluation of the National Institutes of Health Consensus Development Program* (Kanouse et al., 1989).

A set of recommendations for improving the NIH CDP were developed during the meeting and further refined in subsequent committee conference calls and related discussions involving committee members and staff. These recommendations are presented in Chapter 2, Recommendations for Program Structure, and Chapter 3, Recommendations for Program Function. The recommendations are printed in italics and are followed by a summary of the committee's underlying rationale. The committee's deliberations were constrained by the limited time and financial resources made available for this study. Nevertheless, the committee was able to draw upon OMAR's formal documentation of the NIH CDP; the background paper prepared by IOM staff for the committee; previous reports on the NIH CDP; published articles about the NIH CDP, other consensus development programs, group judgment and process, and data integration and synthesis methods; presentations and discussions at the meeting; participants' experiences as NIH CDP panel chairs, members, speakers, consultants, and observers; and participants' experience with other consensus development and group judgment activities and with data evaluation and integration methods used in the health field.

2

Recommendations for Program Structure

PURPOSE AND SCOPE OF THE PROGRAM

Scope of Topics as Defined by Program Purpose

The stated purpose of the NIH CDP should be expanded to include issues related to the management of clinical conditions in addition to the evaluation of specific biomedical technologies.

According to the OMAR pamphlet "Guidelines for the Selection and Management of Consensus Development Conferences," the NIH consensus development conferences are organized to "produce Consensus Statements on important and controversial topics in medicine." However, the current OMAR guidelines proceed to describe a more restrictive purpose for the NIH CDP:

> The purpose of a Consensus Development Conference is to evaluate the available scientific information on a biomedical technology and to produce a Consensus Statement that advances understanding of the technology or issue in question . . . .

In practice, NIH consensus development conferences have addressed the management of clinical conditions, as well as biomedical technologies. Examples of clinical conditions that have been addressed by NIH consensus development conferences include osteoporosis

(1984), neurofibromatosis (1987), and urinary incontinence (1988).[1] Accordingly, it would be appropriate to include reference to management of clinical conditions in the program's statement of purpose.

Scope of Inquiry as Defined by Program Purpose

The scope of inquiry of the NIH CDP should be expanded. The program should seek to ensure that relevant economic, social, and ethical aspects of assessing biomedical technologies and management of clinical problems are appropriately addressed as part of the consensus process. OMAR should commit the necessary resources to identify and, where appropriate, evaluate these aspects. Specifically, OMAR should do one or more of the following for each conference topic: expand or modify the structure of consensus development conferences (i.e., planning committee and conference panel membership, questions, speakers, and consensus statement) to address these aspects of assessment; convene a second panel to address the relevant economic, social, and ethical issues and prepare a statement to be issued in conjunction with the consensus statement on the biomedical aspects of a topic; and request another appropriate organization to prepare a statement that addresses these aspects and to issue that statement in conjunction with the NIH CDP consensus statement.

The scope of inquiry of the NIH CDP should reflect its role in technology assessment[2] and the intent that the program should have an impact on health care practice. As a prominent program of health technology assessment, the NIH CDP has the potential to influence decisions about clinical practice, health technology acquisition and use, insurance coverage and reimbursement, product development,

[1] A complete list of conference topics is provided in Appendix B.

[2] The committee's definition of *technology assessment* is consistent with the definition outlined by IOM; that is, medical technology assessment is "any process of examining and reporting properties of a medical technology used in health care, such as safety, efficacy, feasibility, and indications for use, cost, and cost-effectiveness, as well as social, economic, and ethical consequences, whether intended or unintended" (IOM, 1985).

and public health policies. As noted in the recent RAND Corporation study of the program's impact on medical practice,

> The CDP is more than an assessment program. It is also a communication program to the professional community and the public. It aims to disseminate the results to health care professionals (as well as researchers) throughout the country *in order to improve the state of professional practice*. The expectation is that once physicians and other relevant health care professional personnel know the expert consensus about a particular medical procedure, device, or condition, they will change their practice to conform to the consensually validated recommendation. Increasing public knowledge about the technology under review will help to encourage this change (Kanouse et al., 1989) (emphasis added).

This recommendation regarding the scope of inquiry of the NIH CDP recognizes the overarching mission of NIH.

> Through the conduct, support and promotion of biomedical research, NIH seeks to improve the health of the American people by: increasing the understanding of the processes underlying human health, disability, and disease; advancing knowledge concerning the health effects of interactions between man and the environment; developing methods of preventing, detecting, diagnosing, and treating disease; and disseminating research results for critical review and ultimately for medical application (NIH, 1989).

Assessment of technology and related health practices has the potential to affect people's health by influencing their health habits, health care providers' actions, and resource allocation within the health care system. But, evaluation of safety and efficacy provides only part of the information needed by health care professionals, patients, third-party payers, and other decision makers. To be effective in improving health care practice, health technology assessments must also address relevant economic, social, and ethical consequences, such as cost, access, and quality of life.

By limiting its explicit attention to matters pertaining to evaluation of safety and efficacy, the NIH CDP does not adequately meet the needs of health care professionals, patients, policymakers, third-party payers, industry, and members of the general public who look to NIH consensus statements for authoritative guidance. If NIH is to undertake consensus development, it must seek to ensure that applicable economic, social, and ethical issues are addressed in order to provide useful information about a technology to interested parties and decision makers, even if such analyses extend beyond the core mission of NIH.

Expanding the structure of the NIH consensus development conferences to address relevant economic, social, and ethical aspects has the advantage of providing explicit attention to these aspects. It is important that such matters as cost, accessibility, equity, societal norms, and legal implications be considered in a shared context with technologic performance, safety, efficacy, and other strictly biomedical issues in order to deal explicitly with the trade-offs among these that are raised in research settings as well as clinical practice. However, within the current time constraints allotted for conferences, expansion of scope may become too burdensome to produce useful results. Therefore, it may be necessary for OMAR to convene a second panel to produce a statement, or to have a statement prepared by an outside organization, in order to address these aspects adequately. Such a supplemental effort could be performed concurrent with or subsequent to the consensus development conference on issues of safety and efficacy. In some instances, certain salient economic, social, and ethical issues may not become evident until consensus on the strictly biomedical issues has been achieved. In those cases, OMAR could provide for a second panel to address such issues. Convening a second panel or contracting with another organization would require additional coordination by OMAR, and, if conducted in a sequential rather than parallel manner, would add time to the consensus development process.

In particular, it may be appropriate for NIH to coordinate a two-stage or other complementary process with the Agency for Health Care Policy and Research (AHCPR). As provided in Section 904 of P.L. 101–239 (U.S. Congress, 1989), AHCPR "shall conduct and support specific assessments of health care technologies" and in so doing "shall consider the safety, efficacy, and effectiveness, and, as appropriate, the cost-effectiveness, legal, social and ethical implications, and appropriate uses of such technologies, including consideration of geographic factors." Alternatively, other organizations in government or the private sector could be requested to provide assistance to NIH in addressing relevant economic, social, and ethical issues.

Program Purpose

OMAR should expand the purpose of the NIH CDP to acknowledge explicitly that the ultimate goal of the program is to change behavior toward appropriate use of health practices and technology.

A clear and accurate statement of purpose is essential to the success of the CDP. A clearly stated purpose provides guidance to conference planners and participants and assists users of conference findings by delineating what is to be addressed by a conference, the intended target of the resulting information, and the reason for generating this information. Such a statement of purpose should define the program scope by identifying the appropriate range of topics and extent of inquiry to be addressed.

The current purpose of the NIH CDP as described in OMAR's guidelines is

> to evaluate the available scientific information on a biomedical technology and to produce a Consensus Statement that advances understanding of the technology or issue in question (assessment) and that will be useful to health professionals and the public at large (transfer) [*sic*] (OMAR, NIH, 1988).

The statement of purpose should be expanded to encompass behaviorally oriented program objectives. This will assist NIH in developing, conducting, disseminating, and evaluating the CDP. Clearly stating that the purpose of the NIH CDP is to influence behavior regarding the appropriate use of medical interventions will establish a focused program goal or standard toward which the program can be managed and against which its impact can be measured. Specifically, the statement of purpose should explicitly note the program objectives to promote the timely incorporation of beneficial medical innovations into clinical practice, encourage the abandonment of obsolete technologies in favor of ones that are more efficacious or safe, discourage the adoption of technologies that have little value, and inform public policy choices that encourage or discourage the use of certain medical technologies (Kanouse et al., 1989). Each of these objectives seeks to influence behavior in order to improve health care practices, but is not given due weight in the current NIH CDP statement of purpose.

Objectives of Individual Consensus Development Conferences

The objectives of each consensus development conference should be clearly stated as part of the planning process (including topic selection), the conference itself, and the consensus statement. To aid in delineating objectives for a particular conference, OMAR should consider developing a typology that identifies various potential conference objectives and the audiences that the conference might serve.

In fulfilling the mission of the NIH CDP, an individual consensus development conference can meet one or more objectives for one or more audiences. The objectives of an individual conference might include education, policy-making, resolution of controversial issues, or examination of specific clinical practices. The audience for the resulting consensus statement might include researchers, clinicians, third-party payers, health care administrators, professional organization representatives, consumers (patients), or the general public. Such variation in objectives and audiences among consensus development conferences is necessary and appropriate given the diversity of topics to be addressed. Explicit acknowledgment of objectives and intended audiences will assist in conference planning, conduct, and dissemination efforts.

ORGANIZATIONAL ISSUES

OMAR is part of the Office of the Director of NIH, where it reports to the Associate Director for Disease Prevention. OMAR is the focal point in NIH for technology assessment and transfer. The aim of OMAR's activities is to facilitate the transfer of NIH-supported biomedical research results into clinical applications and to evaluate these research findings for safety and efficacy. In addition to the CDP, OMAR coordinates NIH medical and scientific review of Medicare coverage issues,[3] conducts research and evaluation studies of NIH technology assessment and transfer efforts, and serves as liaison between NIH and health professionals and the general public (IOM, 1985). As a result, OMAR has a varied constituency that includes the BIDs of NIH, biomedical researchers, health care providers, professional societies, voluntary organizations, private industry, state and local governments, other federal agencies, consumer groups, and the general public (OMAR, NIH, undated).

Location of OMAR within NIH

Organizational support for the NIH CDP should be strengthened at the highest levels of NIH. Specifically, OMAR's reporting rela

[3] These issues are referred to NIH and other Public Health Service agencies (e.g., the Food and Drug Administration) from the Health Care Financing Administration primarily via the Office of Health Technology Assessment (OHTA). OHTA was formerly part of National Center for Health Services Research and is retained in the Agency for Health Care Policy and Research.

tionship to the director of NIH should be reevaluated with respect to the CDP. Further, OMAR should be an equal partner with the BIDs in setting the agenda of the CDP.

Given NIH's traditional focus on biomedical research, some observers may regard technology assessment and transfer as subordinate or marginal concerns of NIH. Nevertheless, the functions of technology assessment and transfer are essential to the mission of NIH. In order to have an impact on the health of the public, the knowledge gained through the efforts of NIH and other scientific programs must be transmitted to health care providers and others in the health care system. Successful accomplishment of technology assessment and transfer requires that the needs of the intended audiences of these activities are represented and understood. Thus, the scope of the CDP—especially as reflected in the selection of consensus development topics and the drafting of conference questions—must be responsive to health care providers and others in the health care system. These needs should be reflected in the CDP, but may not, in all instances, be consonant with the particular priorities of the respective BIDs of NIH. Furthermore, OMAR's role in technology assessment and transfer, in particular, as coordinator of the CDP, offers opportunities for NIH to gain the insights of representatives of the broader health care system that may be useful to the agency in formulating its research agenda and related policies.

To ensure that the functions of technology assessment and transfer remain integral aspects of NIH activity and that OMAR is able to adequately represent the concerns of those outside of NIH in the CDP, OMAR's organizational role should be strengthened. There appears to be no particular advantage to the current placement of OMAR in the Office of Disease Prevention; indeed, this may have the potential, or give the appearance, of skewing the emphasis of the program. Stronger organizational support would be best achieved through a closer, perhaps direct, reporting relationship between the directors of NIH and OMAR. Further, OMAR should be an equal partner with the BIDs in setting the agenda of the CDP.

OMAR Advisory Council

An external advisory council should be established to assist OMAR in setting its agenda, including provision of oversight to the NIH CDP and guidance in consensus development conference topic se

lection. The council should be broadly constituted so that the views of health care professionals, provider organizations, industry, researchers, third-party payers, and the general public can be heard.

In order to be effective in technology assessment and transfer, OMAR must have an understanding of the perspectives and priorities of the broader health care system. Currently, there is no standing means for providing OMAR with the views of its constituencies outside NIH. Formalizing OMAR's link to the various segments of the health care system would benefit the CDP and may prove useful for OMAR's other technology assessment and transfer activities as well.

PROGRAM EVALUATION

OMAR should develop an explicit ongoing research effort to determine ways to improve the NIH CDP and to monitor the impact of the program.

This report constitutes just one step in a continuing process of improvement of the NIH CDP in particular and of the methodology of consensus development in general. OMAR should experiment with new techniques (see the recommendations in the section Development of Consensus and Group Dynamics in Chapter 3 of this report) to improve its program and to provide guidance to other consensus development efforts. As noted in Chapter 1 of this report, dissemination of CDP findings is as essential an element in the process as the other program components addressed here are, given the intent to use consensus statements to influence health care practice. Thus, OMAR should build on its efforts to improve dissemination by engaging in further analysis of this area. In addition, evaluation by OMAR and outside organizations of the impact of the NIH CDP should receive high priority and should be performed on an ongoing basis.

FINANCIAL SUPPORT

OMAR should revise its budget expectations for consensus development conferences in light of the recommendations of this report, which likely would require additional resources. If NIH is unable to

increase the funding available to OMAR for the consensus development conferences accordingly, then OMAR should reduce the number of conferences it sponsors to have adequate funding for each conference held.

A number of the committee's recommendations will result in an increase in the amount of money spent per conference and in administrative costs for OMAR (e.g., pertaining to the broader scope of the program, use of a professional facilitator, establishment of an OMAR advisory council, meta-analysis, decision modeling, and ongoing evaluation). Given the importance of the NIH CDP and the authority attributed to it, OMAR should focus on the quality, rather than the quantity, of its consensus development conferences. It is more important to conduct fewer conferences that address the full range of issues required by decision makers than to conduct a greater number of conferences that do not provide decision makers with adequate information to change behavior.

3

Recommendations for Program Function

PLANNING/PREPARATION

Topic Selection

The topic selection process should generate and select topics that reflect the priorities of the broader health care system. Specifically, OMAR should have a standing mechanism for actively soliciting topic suggestions throughout the health care system on a regular basis. This solicitation should include, but not be limited to, professional organizations, specialty societies, third-party payers, health care providers, government agencies, and the general public. OMAR should describe the process for generating potential topics to facilitate contributions by interested individuals and organizations. The process of selecting topics should be formalized and include participation by the aforementioned OMAR advisory council, OMAR staff, and BID directors.

The selection of a topic is a critical determinant of the value of a consensus conference. The NIH CDP has been criticized for examining topics that do not merit the evaluation or expenditure associated with the program. Thus, it has been suggested that "there should be a mechanism for setting priorities among the approaches to be assessed" (Perry, 1987). The current process for selecting

topics, as described in the OMAR Guidelines (OMAR, NIH, 1988), relies largely on input from within NIH, without adequate representation of other segments of the health care system or consideration of important nonbiomedical issues.

The active solicitation of topic suggestions on a regular basis will assist OMAR in identifying the concerns of the health care system. Examples of agencies and organizations that might provide topic suggestions include, but are not limited to, the Agency for Health Care Policy and Research, American Academy of Family Practitioners, American College of Physicians, American College of Surgeons, American Hospital Association, American Medical Association, American Nursing Association, Blue Cross and Blue Shield Association, Centers for Disease Control, Council of Medical Specialty Societies and its member organizations, Department of Veterans Affairs, Food and Drug Administration, Health Care Financing Administration, Health Industry Manufacturers Association, Health Insurance Association of America, National Institutes of Health, and Pharmaceutical Manufacturers Association.

A formal process involving an OMAR advisory council, OMAR staff, and BID directors should be used to set priorities among potential topics. Consistent with the earlier discussion on the relationship of consensus development to changing behavior, the criteria for topic selection should be augmented to reflect an emphasis on the potential to improve the effectiveness of clinical practice, as well as the feasibility of achieving such results. For each potential topic, it will be useful to identify the specific groups that most need or are likely to use the consensus findings, the use that such groups might make of the information, and the approximate magnitude of impact that the findings are anticipated to have within a specified time frame.

Evidence on the state of clinical practice associated with potential topics related to existing technologies or practices should be analyzed to help OMAR in selecting topics, the planning committee in drafting questions, and the panel in developing its statement.

In determining whether a potential CDP topic is controversial and whether a conference on that topic is likely to be useful, OMAR should consider the available information about pertinent practice patterns as well as the current state of the applicable science (Kanouse

et al., 1989). The extent to which there is evidence about clinical practice variation, potentially inappropriate utilization, and related aspects of practice should be taken into account when choosing CDP topics. Evidence should be examined on the extent to which practices are used, the circumstances under which they are used, and the use of any particular relevant techniques (e.g., medical or surgical). Clinical practice information will help planning committees to establish the scope of consideration for CDP topics and to target conference questions. Such information may be necessary for determining whether matters of delivery or access, or related economic, social, or ethical issues, should be considered. Panels may use clinical practice evidence for determining, in part, how their recommendations might be expected to affect clinician behavior (e.g., result in the reduction, increase, or other modification in the use of a procedure).

OMAR could review recent data collected for other purposes (e.g., Medicare claims data, data collected by Peer Review Organizations, the National Disease and Therapeutic Index, or the National Ambulatory Medical Care Survey) (Kanouse et al., 1989). OMAR should coordinate this effort with AHCPR, which is to conduct research and evaluation on the extent to which rates of utilization vary among similar populations for particular diseases, disorders, and other health conditions; uncertainties exist on the effect of utilizing a particular service or procedure; or inappropriate services and procedures are provided (U.S. Congress, 1989).

Some medical practices and technologies have been widely adopted, despite the lack of reliable information about benefits and risks (Asch and Lowe, 1984). Even when evidence is not adequate to support firm conclusions about safety and efficacy, NIH consensus development conferences on management of important clinical problems may be useful. For example, a consensus development conference might evaluate an important, frequently used technology for which evidence of safety and efficacy is inadequate for developing recommendations about its appropriate use. Such a conference could identify specific areas for further research and would provide a more tenable basis for health care delivery, payment, and related decisions by informing clinicians, patients, provider organizations, and third-party payers that adequate evidence for well-grounded recommendations on the technology's use is lacking.

The evidence criterion for topic selection should be broadened to include various types of scientifically sound evidence; in addition to evidence on safety and efficacy, such evidence might include practice patterns and utilization, access, and the findings of other relevant health services research.

Currently, for a topic to be selected for consensus review, "the topic must have an adequately defined and available base of scientific information to answer the previously posed questions and to resolve the controversies insofar as possible" (OMAR, NIH, 1988). Given the above recommendations to broaden the scope of the program and to consider evidence on the state of clinical practice in selecting topics, the existing topic selection criterion regarding data needs should be broadened.

Planning Committee Formation

The planning committee members should include members who understand the clinical, methodologic, and societal issues related to the conference topic. The planning committee should include representatives of relevant BIDs, health care providers (physicians, nurses, other health care professionals, and delivery organizations), third-party payers, government agencies, health services researchers, epidemiologists, biostatisticians, economists, ethicists, patients, and the general public, as relevant to the topic at hand.

The planning committee for each conference greatly influences the outcome of a consensus development conference by drafting the questions that define the scope and substance of the conference, by nominating panelists who will respond to those questions, and by nominating the speakers who will present the information on which, in large part, the panelists will base their findings. Accordingly, the membership of the planning committee should reflect a balance of the parties interested in the selected topic in its entirety and possess a comprehensive understanding of the conference topic (as broadly defined above). However, achieving broader membership on the planning committee should not compromise the efficiency of that group, and OMAR should balance the needs for broad representation

and effective group size when appointing planning committee members.

The planning committee chair should be an individual who is not employed by NIH. The planning committee chair should be selected in consultation with OMAR staff, relevant BID directors, and the OMAR advisory council recommended above.

As noted above, the planning committee shapes the consensus development conference and can greatly influence its outcome. There is a high degree of NIH staff member participation in the CDP. It is desirable to complement NIH staff contributions to the process by assigning the role of planning committee chair to an individual who can broaden representation from other health care sectors.

Drafting Conference Questions

The planning committee should publicly solicit questions concerning a selected topic from a broad base of relevant organizations and individuals, including OMAR advisory council members. This could be accomplished through announcements in professional journals, the Federal Register, *and other notices and by contacting professional, provider, industry, government, and patient organizations.*

Conference questions[1] should be "selected by means of a more thorough search procedure" (Wortman and Vinokur, 1982). The questions posed to a CDP panel should reflect a broad spectrum of health care concerns about the conference topic. In order to understand and include the perspectives of the various health care system participants in the questions, representatives of these participants should be consulted during the question preparation process.

Persons not involved in the process of drafting conference questions should review the questions to ensure that they are clearly stated and will be understood by the consensus panel, conference participants, and readers of the consensus statement.

[1] A list of the conference questions from the NIH consensus conferences held between October 1987 and September 1989 is provided in Appendix C.

Conference questions should be stated clearly to eliminate ambiguity or biased interpretation. Careful and precise definition of the questions to be addressed is required to investigate the topic efficiently (Fink et al., 1984), to maximize the opportunity for a valid evaluation, and to avoid a predetermined conclusion (Oliver, 1985). Poorly worded questions can result in ambiguous or overly generalized recommendations (Lomas, 1986).

Speaker Selection

Suggestions for speakers should be solicited from a wide range of knowledgeable persons in NIH, other government agencies, professional organizations, and other appropriate groups.

A more thorough search procedure should be used in selecting speakers (Wortman and Vinokur, 1982). Presentations by conference speakers should reflect the breadth of knowledge and diversity of opinion about a topic. Solicitation of suggestions for conference speakers will enhance the planning committee's ability to determine when differences of opinion exist about a topic and to "include the presentation of opposing data and interpretations" (OMAR, NIH, 1988). Broadening the process for selecting speakers will lessen the potential for biased selection of speakers (Perry, 1987).

Consensus Panel Selection and Composition

The panel chair plays an important role in shaping the outcome of consensus conferences (Wortman and Vinokur, 1982). Thus, as OMAR's guidelines state, the chair should be "knowledgeable and prestigious" in the field of medical science under consideration, but should not be "identified with strong advocacy positions regarding the consensus topics or with research that might be presented" (OMAR, NIH, 1988). Further, the individual selected should have strong leadership and group process skills. There is a risk that a panel chair may become too influential, particularly in light of the panel chair's participation in the planning committee. Panel chair participation in the planning committee provides continuity of the planning and implementation of conferences, enhances the panel chair's understanding of the process and the specific conference objectives, and develops a stronger relationship between the chair

and OMAR staff. However, the objectivity of the process may be compromised by the potential for one individual to exert excessive control over the outcome of the process.

This committee did not reach agreement on whether the panel chair should be a member of the planning committee. Although a majority of committee members indicated that the benefits of involving the panel chair in the planning process outweigh the potential risks, some did not. A number of committee members indicated that OMAR should explore modified roles of the panel chair (e.g., not having the panel chair serve on the planning committee or introducing a professional facilitator to assist the panel chair, as discussed below).

Suggestions for panelists should be solicited from a broad range of organizations and individuals interested in the consensus topic.

Solicitation of suggestions for panel members from both within and outside of NIH will enhance the planning committee's ability to achieve "balanced representation from various sectors of professional and community life" (OMAR, NIH, 1988). Achieving this representation will validate the panel recommendations (Asch and Lowe, 1984). As noted above, formal solicitation of suggestions for conference participants may help to offset the potential for bias in the process.

Conference panelists, when possible, should possess overlapping expertise.

The breadth of panel expertise should be commensurate with the issues being addressed by a conference. When assembling the conference panel, the planning committee should carefully identify and distinguish among the various disciplines required to address conference questions. For example, the related fields of biostatistics and epidemiology may both need to be represented on a conference panel.

Achieving the necessary breadth of expertise must be accomplished while maintaining a manageable group size for the panel (Wortman and Vinokur, 1982). Selection of panelists who possess more than one area of expertise facilitates the balancing of group size and group expertise. Further, there are potential disadvantages associated with representation of a discipline by a single expert. First, the overall composition of the consensus panel can influence the degree

of participation by individual panelists in the consensus development process; a single panel member with a particular expertise might be isolated during panel deliberations. Second, sound decisions are best achieved when the panel is not forced to defer to a single individual on the panel for crucial expertise. The selection of panelists with multiple areas of expertise reduces the likelihood of a panel containing a solitary expert, without increasing the size of the panel.

IMPLEMENTATION

Preparation of Speakers, Data Synthesis, and Data Presentation

Speakers should submit, in advance of the conference, brief statements that outline their positions, reasoning, or findings on each specific question they have been asked to address; the full text of the key referenced articles on which they base their positions; and, preferably, the full text of their presentations.

Panelists must deal with a large amount of evidence during the consensus conference. Prior to the conference, panelists should receive the information to be presented at the conference in a well-organized format. This will provide more time for them to evaluate the information and prepare any questions for the conference speakers. Providing these materials to the panelists before the conference may help to diminish any undue influence of speakers' manners of presentation or personal characteristics on the panels' evaluation of the evidence being presented. Further, this information can serve as reference material to the panel during its deliberations.

Whenever possible (i.e., when the nature of available data makes this feasible), meta-analysis[2] should be performed before the panel

[2] *Meta-analysis* is a statistical method for obtaining quantitative findings, e.g., on the effect size of a medical intervention, from multiple reports of primary studies on a particular subject. From information obtained from each primary source, a synthesis is made that may produce a stronger conclusion than that which any of the separate reports can provide. It is generally most appropriate when there are no definitive studies on a topic and the nondefinitive studies are in some disagreement (Bulpitt, 1988; Hunter, 1982; Louis et al., 1985; Pillemer and Light, 1980; Sacks et al., 1987; Thacker, 1988).

is convened in order to obtain more rigorous interpretations of multiple sources of evidence.

Evidence presented at a consensus development conference may be complex as well as conflicting. Therefore, it is often difficult for panelists to analyze and synthesize the information presented in the short period of time normally provided. Use of data analytic and reduction techniques (e.g., meta-analysis) has resulted in evidence being "well integrated into both the deliberations and the consensus statement" (Jacoby, 1988). A critical element of meta-analysis is that the process takes account of the quality of the information, as well as information about the history, setting, types of patients, and other factors that may explain differences across studies. Other issues to be considered in reviewing the evidence include the precision of definition of the outcome being measured, the adequacy of the study methodology and the degree to which it has been described, the adequacy of sample size, the degree to which the characteristics of the population studied and the activity being evaluated have been described, and the degree to which results can be generalized (Public Health Service, 1989).

Given the subject matter and types of questions posed for a particular topic, more than one meta-analysis may be prepared for a particular consensus development conference. The findings of a meta-analysis conducted on any particular research issue may vary according to such factors as the thoroughness of the literature search for research reports and analysts' decision rules about accepting and grouping reports in a meta-analysis. Meta-analysis alone is unlikely to fully address the major concerns of a consensus conference. It does offer a rigorous means of analyzing and weighing evidence and rendering more explicit areas of uncertainty in available research.

OMAR should employ available statistical and related methodologic techniques to synthesize the evidence into an understandable format that will facilitate the assessment of evidence by the panel. This can be accomplished by contracting with consultants who have expertise with these techniques and some working knowledge of the conference topic, or by developing the necessary staff expertise to perform meta-analysis and to work with representatives of a particular BID who are familiar with the topic. A combination of experts familiar with meta-analysis and with the topic area is likely to be most desirable.

Evidence should be presented to the panel in a standardized format and be graded in a manner consistent with its quality.

The connection between evidence and recommendations has been unsatisfactory for some consensus conferences (Ahrens, 1985; Perry, 1987). As a result, some recommendations from the NIH CDP may have been premature or inadequately supported. The process of evaluating evidence and deriving recommendations from it should be explicit. As noted above, panelists should consider not only the results and findings of the studies but also the strength of the study design and the validity of the findings when developing consensus. Although there are potential disadvantages as well as advantages[3] associated with grading evidence, it can be an excellent means of improving the process of evaluating evidence and deriving well-founded recommendations. For example, a system of grades of evidence similar to the one used by the Canadian Task Force on the Periodic Health Examination and the U.S. Preventive Services Task Force might be appropriate.[4]

- The process is explicit for speakers and panelists, reducing the extent to which confusion about the quality of evidence, bias, and vested interest can influence the recommendations.
- The process is explicit for members of the health care profession and the public, so that the way evidence was handled can be appreciated and challenged. Those who have different standards can apply them to the same evidence. Those who think that important evidence was not considered can protest or incorporate it themselves.
- When new or additional information becomes available, it can be evaluated and, if appropriate, incorporated.

Potential disadvantages include the following.

- Any simple method of grading is necessarily arbitrary. •Additional staff work is required, but this will be offset by the panel being able to direct its energies to the most important evidence.
- Debate exists surrounding how the levels of evidence are to be translated into grades of recommendations and how much evidence (if any) should be cited along with recommendations.
- The standards for determining the strength of evidence will differ somewhat for different technologies (e.g., preventive, therapeutic, and diagnostic).

[3] Potential advantages include the following.

[4] The Canadian Task Force on the Periodic Health Examination and the U.S. Preventive Services Task Force have used the following system for grading evidence.

Presentation of non-peer-reviewed original evidence should be limited to those extraordinary instances in which the conference panel and planning committee chairs determine that compelling reasons exist for its presentation to the panel. Consistent with the preceding recommendation, any such evidence should be clearly graded.

The time available for the convening of panelists in consensus conferences is limited. Staff and panelists therefore should be afforded ample time before a conference to carefully review and weigh the considerable amount of evidence that may be presented for a consensus conference topic. This process will enable panelists to ascertain and compare the quality of studies and other data. Certainly, peer-reviewed evidence has been subject to a measure of expert examination for quality that most panelists otherwise would be unable to perform in the context of a consensus conference session. Expert presentations made at the conferences of peer-reviewed studies that have been examined by panelists prior to the conference can focus on key issues of direct relevance to the conference, and panelists will be better prepared to make useful inquiries.

The apparent drawback of limiting conference presentations to previously available peer-reviewed evidence is that sufficient time may not have been available for very recent studies to have been

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 studies with or without intervention. Dramatic results in uncontrolled experiments (such as the results of the introduction of penicillin in the 1940s) could also be regarded as this type of evidence.

III. Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees (Goldbloom et al., 1989; Harris et al., 1989; Sackett, 1989)

Group III can be further classified into:

III-1. a group opinion with formal integrative technique,

III-2. a group opinion with informal consensus, and

III-3. an individual's opinion.

peer reviewed and published. Even so, except under extraordinary circumstances, evidence of unknown quality should not be introduced to panelists for the first time in the limited format of a conference, where it would be difficult to ascertain how it affects the balance of evidence already at hand. If compelling reasons exist for introducing such evidence, every effort must be made prior to the conference to subject it to grading and other critical review. Further, as final consensus statements should reflect the evidentiary basis of their findings, readers will be better prepared to consider whether the findings of evidence peer reviewed after the conference might alter the consensus statements' findings and whether a reevaluation of the topic is in order.

OMAR's guideline that "speakers should be asked to confine their presentations to the scientific topic that they have agreed to address" should be strictly enforced by the panel chair during the conference. This guideline should be expanded to require speakers to focus on the scientific evidence related to the question being addressed, rather than relying on anecdote or opinion.

As noted above, a relatively short period of time is available during the conference for presentation of evidence, and a great deal of evidence must be considered by the panel. In order to make efficient use of the available time and to avoid burdening the panel with irrelevant evidence, speakers should be limited to appropriate evidence in their presentations. Evidence considered as part of the NIH CDP should be scientifically sound and not based on anecdote or opinion.

Development of Consensus and Group Dynamics

A working definition of consensus should be made available to, and agreed upon by, panels prior to initiating deliberations and should be included in the consensus statement.

There is no standard definition of *consensus*; that is, it may connote unanimity, majority rule, or other levels or expressions of agreement. A panel should specify an operational definition of consensus (e.g., full agreement, majority agreement, voting, or mean score) to

ascertain the level of agreement during the deliberations and to reveal the existence of substantial minority opinions (Fink et al., 1984). The definition of consensus and level of agreement actually achieved should be reflected in the consensus statement to provide an indication of the strength of the recommendations.[5]

OMAR should experiment with quantitative decision modeling as a means of supporting the work of consensus panels.

A panel's charge to reach consensus within the allotted time may be difficult. Although benefits are not yet fully known, OMAR is encouraged to explore the use of different means of consensus development support in order to improve the NIH CDP and provide guidance for other consensus development efforts.

Decision modeling addresses well-focused questions that involve substantial uncertainty in order to "provide a decision maker with an explicit reproducible process that can help structure complex alternatives in a rational way" (McNeil and Pauker, 1984). The basic technique involves developing a formal, explicit model of questions under consideration. Such a model provides language for representing conflicting data and conflicting objectives and can be explored under a variety of alternative assumptions to determine whether or not the choice of an optimal strategy is dependent on variations in the data base (Pauker, 1986).

Quantitative decision aids and their use of explicit processes to assist in evaluating evidence and deriving recommendations may be of value to panels in achieving consensus. Such tools that facilitate achieving "mathematical consensus" might provide a more cost-effective approach to consensus development and avoid the risk of inadequately considered conclusions because of the potentially onerous night sessions. Implementation of more formal definitions of consensus would be facilitated by the use of these methods. Panels should be provided with information and decision tools to support their efforts. The program should experiment with these to deter

[5] One way to indicate the strength of recommendations would be for panels to indicate the certainty of their recommendations qualitatively or quantitatively. The "lowest common denominator" of acceptance is a potential problem for many issues, and might be alleviated by providing a better indication of the degree of consensus actually achieved. Readers may want to know, for example, whether there is a near-unanimous opinion, with little variance, or whether the result reflects an averaging of extremely diverse opinions.

mine how they can be used most appropriately to serve the needs of consensus development (e.g., decision analysts could be included on the panel or as speakers, or OMAR could contract for decision analysis to be performed in support of the conference). However, if quantitative decision aids are to be used, panels must be receptive to their use and careful consideration must be given to how these tools will be incorporated into the process. Further, use of these tools will require considerable interaction between the individuals performing the modeling and the panel members.

OMAR should experiment with various resources (e.g., professional facilitator) and structures (e.g., varying ranges of conference duration) to aid the efforts of panels to achieve a well-founded consensus.

A professional facilitator may assist the panel chair in managing the consensus development process. The experience of working groups indicates that the opportunity for full participation by all members is desirable and is a function of delicate social balances among the group members. As a fellow professional, the chair of the group is part of the panel's social structure and is in competition with the other members for time and influence. The measure of control given to the chair (i.e., to determine who speaks and when) tends to influence group synergy. An expert facilitator may be able to gain more from the different viewpoints and knowledge bases or skills of the various individuals than could be gained otherwise. Specifically, the facilitator might assist the panel chair in ensuring participation by all panelists, preventing domination by strong personalities on the panel, monitoring and allocating panel discussion time, and keeping the panel focused on the questions to be answered.

The limited time available to consensus panels may raise concerns about the quality of the resulting consensus statements by creating pressure for the panels to produce statements before they have an adequate opportunity to review the available evidence and reach well-founded conclusions. OMAR should continue to offer a conference duration of between two and a half and four days to the planning committee (Elliott, 1989).

The panel should take full account of alternative interpretations or recommendations during the consensus development process. The consensus statement should include well-reasoned minority opinions

when they exist, note when such minority opinions do not exist, identify items that lack adequate evidence for a judgment to be made (i.e., explicitly address what is not known about a topic), and identify specific areas for further research.

A repeated criticism of the program has been that minority views are submerged or obscured (Perry, 1987). According to OMAR, there have only been two consensus statements that have included minority opinions. Full consensus on all issues cannot always be achieved, and valuable information can be produced from a conference that does not reach definitive conclusions.

Specific techniques should be considered to ensure that the panel does not override alternative interpretations of evidence or recommendations. For example, one panel member could be assigned to critique the group process at intermediate points or to play the role of "devil's advocate" (Janis, 1971). Consensus should not be forced or reported if none exists, and consensus statements should indicate the degree of agreement achieved.

Approaches for improving group dynamics and development of consensus are being developed and tested in many fields (Dalkey, 1969; Delbecq et al., 1975; IOM, 1985; Linstone and Turoff, 1975; Olsen, 1982; Policy Research Incorporated, 1977; Porter et al., 1980). OMAR can take a leadership role by looking within and beyond the health field for alternative techniques and models used to engage groups of experts in rendering well-founded and informed findings.

Initial and Final Consensus Statement Preparation

The panel should use the collected background material for the conference as the starting point of its discussion, rather than a draft consensus statement prepared by panel members.

It is inappropriate to have a draft consensus statement before a conference begins,[6] as this may imply prejudgment of an issue that

[6] A previous report to NIH recommended that "panelists should be assigned to questions and write first drafts of answers prior to the conference" or that the panel should meet before the conference and "write an advanced draft of the consensus statement" (Wortman and Vinokur, 1982). The recommendations in that report were intended to reduce the burden on the panel during its executive session, when the consensus statement is written, and to provide the panel with additional time to produce an improved final consensus statement.

has not been fully evaluated. Such a statement may limit panel deliberations for or against particular positions rather than allow more open discussion. However, it is appropriate to present information (e.g., the speakers' positions in writing) to panelists prior to the conference, (see the recommendations in the section Preparation of Speakers, Data Synthesis, and Data Presentation earlier in this chapter). It also may be useful to describe a technology or practice and its uses in an objective background paper prepared for the panel. This paper could provide the basis for a preamble to their consensus statement and the descriptive narrative to which the panel may refer in developing the consensus statement. Moreover, this information should be available to panelists in electronic (i.e., machine-readable) form.

The panel should be provided with all practical resources to facilitate writing the consensus statement. Specifically, OMAR should provide the best current computer support and audio-visual technology (e.g., word processing equipment that includes large screens or projection capability) during the drafting of consensus statements and should consider employing a writer to support panel working sessions and to assist the panel in ensuring that the final consensus statement is presented clearly.

The executive session during which the panel writes the consensus statement has been described as "grueling" (Mullan and Jacoby, 1985). Given the existing time constraint and the potential difficulty of drafting consensus statements, the panel should be provided with appropriate tools to lessen the burden of their work and reduce concern that a deadline forces panels into inadequately considered judgments.

References

Ahrens, E.H., Jr. 1985. The diet-heart question in 1985: Has it really been settled? The Lancet i:1085–1087.

Andreasen, P.B. 1988. Consensus conferences in different countries: Aims and perspectives. International Journal of Technology Assessment in Health Care 4:305–308.

Asch, S.M., and C.U. Lowe. 1984. The consensus development program: Theory, process, and critique. Knowledge: Creation, Diffusion, Utilization 5(3):369–385.

Bulpitt, C.J. 1988. Meta-analysis. The Lancet i:93–94.

Calltorp, J. 1988. Consensus development conferences in Sweden: Effects on health policy and administration. International Journal of Technology Assessment in Health Care 4:75–88.

Dalkey, N.C. 1969. The Delphi Method: An Experimental Study of Group Opinion. Santa Monica, Calif.: The RAND Corporation.

Delbecq, A., A.H. Van de Ven, and D.H. Gustafson. 1975. Group Techniques for Program Planning. Glenview, Ill.: Scott, Foresman.

Elliott, J. 1989. Office of Medical Applications of Research, communication to committee.

Fink, A., J. Kosecoff, M. Chassin, and R.H. Brook. 1984. Consensus methods: Characteristics and guidelines for use. The American Journal of Public Health 74(9):979–983.

Gleicher, N. 1984. Cesarean section rates in the United States: The short-term failure of the national consensus development conference in 1980. Journal of the American Medical Association 252(23):3273–3276.

Goldbloom, R., R.N. Battista, and J. Haggerty. 1989. The periodic health examination. Canadian Medical Association Journal 141:205–208.

Goodman, C., ed. 1988. Medical Technology Assessment Directory. Washington, D.C.: National Academy Press.

Harris, S.S., C.J. Caspersen, G.H. Defriese, and E.H. Estes. 1989. U.S. preventive services task force. Journal of the American Medical Association 261:3588–3589.

Hunter, J. 1982. Meta-analysis: Cumulating Research Findings Across Studies. Beverly Hills, Calif.: Sage Publications, Inc.

Institute of Medicine. 1985. Assessing Medical Technologies. Washington, D.C.: National Academy Press.

Jacoby, I. 1985. The consensus development program of the National Institutes of Health: Current practices and historical perspectives. International Journal of Technology Assessment in Health Care 1(2):420–432.

Jacoby, I. 1988. Evidence and consensus. Journal of the American Medical Association 259 (20):3039.

Janis, I.L. 1971. Groupthink. Psychology Today (November).

Kanouse, D.E., R.H. Brook, J.D. Winkler, J. Kosecoff, S.H. Berry, G.M. Carter, J.P. Kahan, L. McCloskey, W.H. Rogers, C.M. Winslow, G.M. Anderson, L. Brodsley, A. Fink, and L. Meredith. 1989. Changing Medical Practice Through Technology Assessment: An Evaluation of the National Institutes of Health Consensus Development Program. Santa Monica, Calif.: The RAND Corporation.

Linstone, H., and M. Turoff, eds. 1975. The Delphi Method: Techniques and Applications. Reading, Mass: Addison-Wesley.

Lomas, J. 1986. The consensus process and evidence dissemination. Canadian Medical Association Journal 134:1340–1341.

Louis, T.A., H.V. Fineberg, and F. Mosteller. 1985. Findings for public health from meta-analyses. Annual Review of Public Health 6:1–20.

McNeil, B.J., and S.G. Pauker. 1984. Decision analysis for public health: Principles and illustrations. Annual Review of Public Health 5:135–161.

Mullan, F., and I. Jacoby. 1985. The town meeting for technology. Journal of the American Medical Association 254(8):1068–1072.

National Institutes of Health. 1989. Fiscal Year 1991 Justification of Appropriation Estimates for Committee on Appropriations: National Institutes of Health. Volume III. Washington, D.C.: U.S. Government Printing Office. Publication No. 1989-252-073.

Office of Medical Applications of Research, National Institutes of Health. Undated. Office of Medical Applications of Research. Bethesda, Md. Pamphlet.

Office of Medical Applications of Research, National Institutes of Health. 1988. Guidelines for the Selection and Management of Consensus Development Conferences. Bethesda, Md. Pamphlet.

Oliver, M.F. 1985. Consensus or nonsensus conferences on coronary heart disease. The Lancet i:1087–1089.

Olsen, S.A. 1982. Group Planning and Problem Solving Methods in Engineering Management. New York: John Wiley & Sons.

Pauker, S. 1986. Decision analysis as a synthetic tool for achieving consensus in technological assessment. International Journal of Technology Assessment in Health Care 2(1):83–97.

Perry, S. 1987. The NIH consensus development program: A decade later. New England Journal of Medicine 317:485–488.

Pillemer, D.B. and R.J. Light. 1980. Synthesizing outcomes: How to use research evidence from many sources. Harvard Educational Review 50(2):176–195.

Policy Research Incorporated. 1977. A Comprehensive Study of the Ethical, Legal, and Social Implications of Advances in Biomedical and Behavioral Research and Technology: Summary of the Final Report. Baltimore.

Porter, A.L., F.A. Rossini, S.R. Carpenter, and A.T. Roper. 1980. A Guidebook for Technology Assessment and Impact Analysis. New York: North-Holland.

Public Health Service. 1989. Caring for Our Future: The Content of Prenatal Care. Report of the Public Health Service Expert Panel on the Content of Prenatal Care. Washington, D.C: U.S. Department of Health and Human Services.

Sackett, D.L. 1989. Rules of evidence and clinical recommendations on the use of antithrombotic agents. Chest 95:2S–4S.

Sacks, H.S., J. Berrier, D. Reitman, V.A. Ancona-Berk, and T.C. Chalmers. 1987. Meta-analyses of randomized controlled trials. New England Journal of Medicine 316(8):450–455.

Thacker, S.B. 1988. Meta-analysis: A quantitative approach to research integration. Journal of the American Medical Association 259(11):1685–1689.

U.S. Congress. 1989. P.L. 101–239. The Omnibus Budget Reconciliation Act of 1989.

Wortman, P.M., and A. Vinokur. 1982. Evaluation of the National Institutes of Health Consensus Development Process, Phase I: Final Report. Ann Arbor: The Institute for Social Research, The University of Michigan.

Appendixes

A

Background Paper for the Committee to Improve the NIH Consensus Development Program*

Introduction

The purpose of this paper is to offer the Committee to Improve the National Institutes of Health Consensus Development Program uniform background information on the National Institutes of Health (NIH) Consensus Development Program (CDP). The committee's report represents one component of a three-part examination of group judgment methods for assessing medical technologies being conducted by the Council on Health Care Technology.

The following synthesis of available literature about the program is intended to provide a common base of understanding for the examination of the NIH CDP. Following a brief history of the NIH CDP, a general overview of the planning and implementation of consensus development conferences is provided. The paper then reviews the available literature concerning selected major components of the CDP. Each component is addressed in the order that the NIH CDP is conducted.

* This paper was prepared by Sharon R. Baratz, Council on Health Care Technology, Institute of Medicine.

Origin of the NIH Consensus Development Program

Throughout the 1970s, the acceleration of technological innovation in medicine, accompanied by rising costs and increased concerns for the quality of care, generated extensive interest in technology assessment. Donald Fredrickson, then director of NIH, thought that the biomedical research community had a significant role to play in the evaluation of new and existing health care technologies. In October of 1978, the NIH formally established the Office of Medical Applications of Research (OMAR) to act as a link between the medical research community, clinical physicians, and the public.

The CDP was established as one component of the technology assessment and information-sharing activities of NIH (Perry and Kalberer, 1980). Although variations have occurred, the following is a description of the general format of the CDP as it has typically been conducted in recent years.

The Consensus Development Program

The CDP provides a forum for concerned individuals to evaluate medical technologies. Observers have noted that the three-day consensus development conference borrows processes from the scientific meeting, the judicial process, and the town meeting (Jacoby and Rose, 1986; Mullan and Jacoby, 1985). In a consensus development conference, the panelists represent a jury who give the verdict after hearing the expert speakers give scientific testimony concerning the technological case at hand. The presentations by speakers are public, and members of the audience may comment and participate, similar to the proceedings of a town meeting.

The entire planning and implementation process for a particular conference usually occurs over a period of 12 to 15 months, although the planning process has occurred over a period of as short as 5 months (Elliott, 1989). Each conference normally lasts over a period of three days.

The original intention of the NIH CDP was to evaluate emerging technologies with potentially significant health impacts. Nevertheless, even in the early years, panels have addressed technologies already in widespread use that had not been previously scrutinized for safety and effectiveness (Perry and Kalberer, 1980). At present,

consensus development conferences investigate technologies that are new, established, or in widespread use, including some that may be obsolescent. Meetings have been held to reassess topics; for example, adjuvant chemotherapy for breast cancer, originally addressed in 1980, was reexamined in 1985 (Goodman, 1988).

The topics of consensus development conferences have varied substantially since the inception of the program. Conferences tend to focus on a technology, e.g., electroconvulsive therapy (1985) or magnetic resonance imaging (1987), or on a particular health problem and the alternative technologies applied for diagnosis, treatment, or rehabilitation of these, e.g., travelers' diarrhea (1985) or adult urinary incontinence (1988).

In general, the bureaus, institutes, and divisions (BIDs) of NIH suggest consensus topics to OMAR for consideration. Topics are considered from other sources, including other Public Health Service agencies such as the Food and Drug Administration, the U.S. Congress, or organizations outside of government (Goodman, 1988).

OMAR uses the following criteria for the selection of conference topics:

- The subject under consideration should have public health importance. The topic should affect or have broad application to a significant number of people.
- There should be controversy surrounding biomedical/scientific aspects of the topic that would be clarified by the consensus approach or a gap between current knowledge and practice that a consensus development conference might help to narrow.
- The topic must have an adequately defined and available base of scientific information to answer the previously posed questions and to resolve the controversies insofar as possible.
- The topic should be amenable to clarification on technical grounds, and the outcome should not depend mainly on the impressions or value judgments of panelists.

Additional elements desirable for positive consideration of a consensus topic include health care cost impact, preventive impact, and public interest (OMAR, NIH, 1988).

The topic selection process may take from two months to a year or more. A representative from the NIH Coordinating Committee on Assessment and Transfer of Technology (CCATT) usually participates in the initial evaluation of topics for consensus confer

ences. The CCATT member is usually a BID senior staff person appointed by the BID director to "serve as an institutional spokesperson for the BID on issues that relate to the Committee. In addition to being a member of the Committee, each BID representative shall serve as the primary focal point in his or her BID for communication with OMAR and with other NIH units on all matters related to technology assessment and transfer" (NIH, 1983). In the event that a topic under consideration is not suitable for a consensus development conference, OMAR and the BID may choose to conduct a workshop or host a scientific meeting to discuss the topic (Elliott, 1989).

Once a topic has been suggested for a consensus development conference, the sponsoring BID nominates a BID coordinator who is knowledgeable in the area of science under consideration. The responsibilities of the BID coordinator include chairing the planning meeting and representing the BID in managing the conference. A senior OMAR staff person is selected as the OMAR coordinator to work with the BID coordinator and other BID representatives in organizing the conference (OMAR, NIH, 1988). OMAR's focus is on the consensus process, while the initiating BID's contribution concerns the scientific information required for the conference topic (Elliott, 1989). The OMAR and BID coordinators, together with other appropriate BID and OMAR staff, meet to determine the topic's acceptability using the criteria listed above for a consensus development conference. After OMAR and a BID have agreed to act as lead sponsors of a particular consensus development conference, other BIDs and other agencies may agree to cosponsor the event (OMAR, NIH, 1988).

The OMAR coordinator, the BID coordinator, and representatives from other sponsoring organizations meet to discuss the potential scope and date of the conference. The conference chairperson is selected at this time (OMAR, NIH, 1988). The sponsoring BID and OMAR jointly select the chairperson (Elliott, 1989). The chairperson of the consensus development conference is chosen at an early stage in conference planning so that he or she can participate as a member of the planning committee and thus bridge the gap between conference planning and implementation (Elliott, 1989). The planning committee is composed of NIH staff, including OMAR and BID staff, and outside experts from the research community who are not federal employees (OMAR, NIH, 1988).

The planning committee has four major functions: "(1) to draft consensus questions, (2) to draft the conference program, (3) to recommend conference speakers, and (4) to recommend consensus panel members" (OMAR, NIH, 1988). The consensus questions asked vary in number and type; in most instances, panels have addressed from four to six questions. The 1988 OMAR guidelines describe the drafting of questions to be addressed as follows:

A. The agenda of a [consensus development conference] is structured around key questions posed to the panel that serve to determine the scope and substance of the conference.
B. Questions should be structured so that answers can be derived from scientific information and data presented by the speakers. The questions should not be phrased in a manner that requires responses dependent solely upon the subjective judgments or opinions of the panelists.
C. Questions should be straightforward, concise, and constructed so that it will be evident whether consensus has been achieved.
D. Ordinarily, four to six questions are posed, including questions on efficacy, risks, clinical applications, and a final one on directions for future research (OMAR, NIH, 1988).

The planning committee plans the agenda. OMAR hires a support contractor to organize all conference logistics.

The planning committee is also responsible for nominating the panel. OMAR conducts a literature search for all suggested panelists, including the chairperson, to ensure that individuals have no published positions concerning the conference topic (Elliott, 1989). OMAR currently assembles panels with balanced representation comprising individuals with various areas of expertise in order that the panel may better deal with the diverse material presented and that the panel's credibility may be enhanced (OMAR, NIH, 1988). In selecting panelists, OMAR seeks individuals who are "thoughtful, able to weigh evidence, and capable of collaborative work" and who "have no vested interest in the technology being reviewed" (OMAR, NIH, 1988). According to OMAR, panels should include individuals involved in research in the field; health professionals who are users of the technology; methodologists or evaluators such as epidemiologists or biostatisticians; and public representatives such as ethicists, lawyers, theologians, economists, public interest group or voluntary

health association representatives, consumers, and patients. Panelists should be residents of the United States and should not be federal employees, to avoid the appearance of undue federal influence (OMAR, NIH, 1988).

Speakers are experts on the chosen topic and present evidence, as appropriate, to the panel on the safety, efficacy, effectiveness, and service requirements of the technology in question (Goodman, 1988). OMAR guidelines describe the speaker's role as follows:

A. Speakers should be selected for their scientific expertise and may include both clinical investigators and basic scientists as well as general authorities in the field. Where differences of scientific opinion exist, care should be exercised to include the presentation of opposing data and interpretations.
B. Speakers should be asked to confine their presentations to the scientific topic that they have agreed to address and to be certain to present all relevant data and information.
C. To prevent the appearance of bias, the planning committee is encouraged not to include any of its members as speakers unless other experts are unavailable (OMAR, NIH, 1988).

At least one month in advance of the actual conference date, panelists receive abstracts of the speakers' presentations to prepare them for the consensus development conference. Speakers are also asked to bring photocopies of their slides to the consensus development conference for panelists. The BID coordinator is responsible for the supply of overview articles and other supplemental materials for the panelists prior to the consensus development conference. The BID coordinator meets with National Library of Medicine (NLM) staff in order to direct and detail the strategy for preconference information retrieval. Panelists receive a copy of this literature search approximately three months prior to the actual conference date. The BID coordinator determines the criteria for the search, for example, how far back in years NLM staff should search and with which key words (Elliott, 1989). Speakers and members of the audience receive conference materials, which include only speakers' presentation summaries, a conference agenda, and logistical information, at the first public meeting of the panel.

The panel chair divides the panel into subcommittees and designates subcommittee chairs for each question several months before the conference. Each group is asked to prepare a proposed format to

answer one question, along with a draft outline of elements to be considered in answering the question. OMAR requests that this material not be in narrative form. Panelists use the information they receive from OMAR to complete this task and submit the documents to the panel chair at the first panel meeting, which is held the day before the opening of the conference. This question assignment process allows panelists to specialize and focus on the relevant issues presented by particular speakers.

During the preconference meeting, the panelists discuss the consensus development process and outline the consensus statement. Panel members are responsible for writing the specific portion of the consensus statement that corresponds to the questions they were assigned prior to the conference. All panelists remain responsible for the statement as a whole, however, and thus should follow all presentations and deliberations in the conference.

During the next two days panelists hear evidence from experts on the consensus topic and prepare several drafts of the consensus statement. On the morning of the first day, the director of OMAR delivers the charge to the panel. The chairperson of the meeting opens the session and is responsible for conducting the presentations in an orderly and timely manner. Each speaker presents his or her expert knowledge on the consensus conference topic in a predetermined amount of time, usually between 15 and 30 minutes. In general, for every hour of presentations there is a discussion period of approximately 30 minutes. The panelists are given the opportunity to ask questions for further information or for clarifications. Members of the conference audience (which have ranged from 200 to 1,100 individuals, depending on the topic) and the attending speakers may then question the presenters and offer additional information. The chairperson seeks to ensure that the group follows the time schedule and does not digress from the consensus topic or questions.

The public testimony from experts continues for one and a half days. After the first day's presentations, the panelists adjourn to draft the consensus statement. After the second day of presentations (which is a half day), the panel again convenes in executive session to continue drafting the statement. At that time, the OMAR and BID coordinators are present to act as resources for the panel (e.g., to check statistics) and to ensure that the panel adheres to OMAR guidelines regarding the general content of the consensus statement (e.g., to ensure that direct recommendations for specific individuals

or agencies are not included in the statement) (Elliott, 1989). A representative from the Food and Drug Administration is usually available to answer the panel's questions on the regulatory status of the technology. The panel remains in closed session the evening of the second day of the conference until consensus is reached and the statement is drafted. These sessions have continued into the early-morning hours of the third day.

On the morning of the third day, the chairperson reads the draft consensus statement to the speakers and interested members of the public. Members of the audience may make comments or suggestions for changes in the statement. The panel adjourns again after the morning session to discuss the proposed alterations and draws up a final draft of the consensus statement. Following this, the panel meets with the press in the afternoon to relate the group's conclusions. Panelists have approximately two weeks to contact OMAR to offer additional changes. For the most part, these later alterations have tended to be stylistic rather than substantive.

Consensus development conferences sponsored by the National Institute of Child Health and Human Development (NICHD) vary from the format described above in several ways. Topics are reviewed and analyzed by the NICHD Office of Planning and Evaluation according to OMAR criteria for consensus conference topic selection and the topic's relevance to the NICHD mission and research priorities. The planning committee for an NICHD/OMAR-sponsored conference chooses the chairperson, specifies the types of individuals for the panel, and identifies conference questions. The planning committee suggests the individuals who should present evidence as speakers. The final decision on conference speakers involves the planning committee, the panel, and NICHD and OMAR staff. To prepare for the first panel meeting, panelists receive materials from an extensive literature search done by NLM. The first meeting is held 12 to 18 months prior to the consensus development conference so that the panel has time to develop a draft report on the chosen topic, including a list of important issues that need to be discussed during the meeting. The draft report documents the panel's analysis of the literature and includes complete references and an appendix of all materials presented at the conference. The draft report is circulated widely among interested organizations and individuals, who are encouraged to submit written comments and/or give oral testimony at the consensus development conference. The public

consensus development conference consists of presentation by the panel of a summary of the draft report, invited comments, presentation of additional data not previously considered, testimony from individuals and organizations, and a closed executive session for preparation of the consensus statement. After the conference the draft report is modified to coincide with the consensus statement and issued as a final report. The products of a NICHD/OMAR consensus development conference are a monograph and a consensus statement. The approximate cost per NICHD/OMAR consensus development conference is $275,000, whereas for OMAR consensus development conferences this figure is $116,000 (Elliott, 1989).

Review of the Major Components of the Consensus Development Program

The structure of OMAR's CDP comprises the following steps:

- topic selection
- formation of the planning committee and selection of the chairperson
- selection of specific conference questions, the consensus panel, and conference speakers
- presentation and synthesis of conference data
- development of consensus
- preparation of initial and final consensus statements
- dissemination of consensus information

Many participants and observers have examined, commented upon, and made suggestions regarding improvement of the program. The substance of such reviews is summarized here, and is arranged by the components identified above.

Topic Selection

The effectiveness of the NIH CDP may be linked to the selection of appropriate topics at the best time for assessment. A study of physicians' knowledge, attitudes, and practices as a result of eight early conferences (1979–1980) indicated that the choice of topics affects OMAR's ability to accomplish its information transfer goal. Kanouse et al. (1989) contended that the 1980 conference on cesarean sections was successful in reaching physicians, because it ad

dressed "an aspect of care perceived to be problematic" and "performed a unique function in drawing together the accumulated knowledge on this topic and pointing out its practice implications." Other consensus development conferences were found to be less successful in reaching physicians. The researchers recommended that NIH "select topics only after a systematic review of the conditions under which the procedure to be covered by the conference is performed, including a determination as to whether physicians' and hospitals' practice presently conform with the recommendations that a panel is likely to make" (Kanouse et al., 1989).

Political interests or pressures may have a substantial impact on the placement of a subject on the CDP agenda. Individuals involved in the Reye's syndrome conference reported that congressional concern in this area accelerated the adoption of the subject for a consensus development conference (Wortman et al., 1988). The 1983 conference on liver transplantation has been cited as an instance of a politically, rather than scientifically, prompted conference (Markle and Chubin, 1987). According to Markle and Chubin, the emotional aspect of a costly liver transplantation, especially when conducted for children, prompted extensive media coverage and congressional pressure to hold a consensus development conference. Markle and Chubin held that the staff of the National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases and OMAR contended that the procedure was still experimental, given the absence of clinical data to substantiate stable positive outcomes after treatment. Liver transplantation was an experimental technology (i.e., the technology had not yet been adequately evaluated for safety and effectiveness in scientific studies) and not an emerging technology (i.e., the procedure had already passed through clinical trials, but was not yet widely adopted by clinicians). Markle and Chubin (1987) reported the following:

> In early 1980 Charles Lowe, then acting associate director of OMAR, requested that the National Institute of Arthritis, Metabolism, and Digestive Diseases (NIAMMD) supply him with an objective summary of the current state of the art of liver transplantation . . . . On April 7, 1980, NIAMMD staff responded to Lowe's request, providing a state-of-the-art assessment and concluding that "despite recent improvements . . . hepatic transplantation is still largely an experimental procedure with relatively unpredictable outcome in individual cases."

Nonetheless, the conference was held in June of 1983. At issue is the conflict between not selecting topics for which insufficient data

are available for reaching scientifically valid conclusions and the pressure to hold conferences on controversial issues (Institute of Medicine, 1985).

Formation of the Planning Committee and Selection of the Chairperson

Individuals, including representatives from the sponsoring agencies, the conference chairperson, and two or three experts from outside NIH, are invited by NIH staff to join the planning committee. The program planning committee is responsible for the definition of conference questions, formulation of a conference agenda, and nomination of the speakers and panelists. The consensus development conference chairperson is appointed prior to the first meeting of the planning committee. According to OMAR, the chairperson of the consensus development conference is selected for his or her stature as a distinguished physician and scientist, nonadvocacy regarding the conference issues, and personal skills in chairing the open symposium portion of the conference and in leading the consensus panel (OMAR, NIH, 1988).

Selection of the Specific Conference Questions, The Consensus Panel, and Conference Speakers

For the most part, the conference questions posed by the planning committee have been found to be straightforward and concise. A study by Wortman and Vinokur (1982) concluded that the questions tend to reflect the concerns of NIH staff. In their study, Wortman and Vinokur (1982) classified 12 sets of questions from conferences held from 1980 to 1982 by content and found that the majority addressed the benefits and appropriate conditions for use of the technology. Social issues were infrequently addressed; only 5 of 84 questions concerned economic, legal, and ethical inquiries. This observation may be consistent with the stated guidelines of the NIH consensus development program, which indicate that

> The purpose of a [consensus development conference] is to evaluate the available scientific information on a biomedical technology and to produce a Consensus Statement that advances understanding of the technology or issue in question (assessment) and that will be useful to health professionals and the public at large (transfer) . . . . Although many aspects of a technology under evaluation may be discussed, the primary focus of a [consensus development conference] is the technology's clinical applications (OMAR, NIH, 1988).

Controversial questions may be excluded from the conference agenda. Wortman et al. (1988) noted that the association between Reye's syndrome and aspirin was not a question posed for the consensus development conference for Reye's syndrome:

> A decision to eliminate a question that did not have a scientific answer at that time was acceptable to the institute planners involved, but could not be kept entirely out of the conference and on the basis of audience reactions and the concern of some panelists, a brief section had to be added to the consensus statement regarding the issue. Within a year, another U.S. Public Health Service agency, the Centers for Disease Control, made a fairly definitive statement on the issue.

Markle and Chubin (1987) pointed out that the 1983 liver transplantation conference explicitly addressed only technical aspects of the technology, although important social, ethical, and economic issues shaped the conference deliberations:

> If liver transplantation was deemed an experimental procedure by the consensus panel, then third-party payment was unlikely. However, if it was classified as a therapy with all the trappings of success, i.e.,, impressive survival rates and enhanced quality of life, then a government subsidy would be warranted. This emerged as a pivotal issue, yet was barely addressed at the conference. The result was that these issues became implicitly important, shaping the consensus statement in ways unrecognized, or at least not admitted by the participants themselves.

The consensus statement on liver transplantation does not address reimbursement for this very expensive procedure; the existing and potential demand for the procedure; or how such demand could be met in terms of available transplant teams, facilities, and donor organs (Institute of Medicine, 1985).

A broader, more coordinative role to encompass wider social, legal, ethical, and economic assessment concerns in the U.S. Department of Health and Human Services was envisioned for another agency. Perry and Kalberer (1980), in a review of the first two years of the NIH CDP, noted that assessments of technologies occurred at two levels in the U.S. Department of Health and Human Services:

> The NIH, representing interest and expertise in the biomedical sciences, has emphasized technical consensus development—the assessment of scientific and medical aspects of the technology in question. Other agencies in this department also engage in evaluation activities on a smaller scale. The recently established National Center for Health Care Technology [NCHCT] is responsible for coordinating all the assessment activities in the department and has a mandate to concern itself with economic, legal, and ethical implications as well as scientific and medical issues.

According to the former director of NCHCT, several features distinguished the NCHCT from the NIH's CDP:

1. the NCHCT's role included the establishment of priorities for assessment among technologies;
2. the NCHCT was able to fund new areas of research identified and fund technology assessment activities;
3. assessments sponsored by the NCHCT were broad in scope, addressing economic and ethical concerns and the direct impact of the assessments on coverage by Medicare (Perry, 1982).

In 1981, budget cuts arrested NCHCT's activities. Markle and Chubin (1987) inferred that the NIH CDP "assumes that a strict separation of factual and value issues is possible, and further, that objective evidence compels experts to converge on the 'correct' decision."

In contrast to concerns regarding the circumscribed scope of the program, some observers and participants have voiced frustration with the broad agenda of certain consensus development conferences. Four to six questions are usually addressed, and one of these normally pertains to suggestions regarding further research (Jacoby, 1985). A panelist at the conference on the health implications of obesity charged that the group attempted to answer too many questions (Henig, 1985). Henig (1985) indicates that the effectiveness of the consensus development conference may be limited if the planning committee selects a wide scope of questions. Researchers at the RAND Corporation recommended that "consensus studies should focus on carefully defined problems that can be investigated in a timely and economical way" (Fink et al., 1984). The need to clearly define conference questions and focus the conference effort, on one hand, and the desire by some health professionals and members of the public to cover social, ethical, and legal implications of the application of a technology, on the other hand, raise the importance of appropriate attention being given to matters of scope and concern at all conferences.

In addition to selection of conference questions, the planning committee nominates individuals to the panel. Over the course of the program, OMAR has assembled two general types of panels: the "balanced panel" and the "neutral panel" (Jacoby, 1985). The balanced panel is made up of representatives of different points of view on the technology in question. In this case, adversarial opinions are intended to carry the same weight, with equal numbers of advocates on each side. The neutral panel consists of experts with no pub

lished opinion on the topic. In this case, individuals are presumed to be less committed to a particular position and thus, perhaps, more willing and able to consider the evidence with less bias. In recent years, the neutral panel has been used as discussions by balanced panels have been reported to tend to "degenerate into unconstructive debate" (Jacoby, 1985).

Asch and Lowe (1984) recommended that the panel consist of representatives of their respective professions who are experts and thus would validate panel recommendations. The size of panels depends upon the complexity of the problem presented, the number of diverse points of view and specialties required, and the resources available for the actual consensus development conference (Fink et al., 1984). In recent years, panels have been neutral and therefore do not reflect diverse points of view, and the size of the panel has not been limited by available resources (Elliott, 1989).

According to Wortman et al. (1988), panelists should be able to weigh evidence and to write clearly and concisely. The panel begins to write the consensus statement during the one and a half days of presentations by the speakers. Given this limited amount of time for writing consensus statements, writing skills are vital to the effectiveness of the panel (Wortman et al., 1988). Although the credibility of consensus statements may depend upon the formation of apparently unbiased panels, critics of the consensus statement have protested that a panel is "stacked in favor of the clinical status quo or NIH-supported recommendations" (Henig, 1985).

Conference speakers are selected on the basis of their expertise concerning the topic in question and may advocate particular positions. Consumer representatives, lawyers, economists, and ethicists are not always represented at consensus development conferences. In general, the planning committee considers the relevance of these types of experts for a particular conference, given the limited speaking time available. An economist gave a presentation at the consensus conference on urinary incontinence (1988) in order to demonstrate the financial impact of the disability on the health care system.

A report on the consensus development program noted that conference participants are drawn from a fairly narrow segment of the biomedical community and that panelists and speakers are often acquaintances of NIH staff members (Wortman et al., 1988). Individuals rarely participate in more than one consensus development

conference (Elliott, 1989). The panelists and speakers may all be drawn from among individuals who are closely affiliated with NIH; thus, they may tend to be involved in biomedical science rather than in policy or social science areas. NIH staff may have less contact and knowledge of the appropriate individuals outside of the medical research and practice communities. A former director of OMAR recommended a broader and more systematic search for conference participants (Perry, 1987).

A former OMAR officer suggested that "a 'separation of powers' be maintained among the triad of forces [planning committee members, speakers, and panelists] responsible for the final consensus product." Jacoby (1985) held that this division of responsibilities among members of the panel, speakers, and the planning committee, throughout the consensus development conference program, would help ensure an "impartiality of thought and equality of influence." At present, members of the planning committee may participate in the consensus development conference as expert speakers but not as panelists, although this is discouraged by OMAR (Elliott, 1989).

Presentation and Synthesis of Conference Data

Asch and Lowe (1984) noted that the validity of the CDP "capitalizes on the ability and experience of NIH in mobilizing expertise, and in providing the data necessary for the formulation of authoritative statements." The panelists receive information on the consensus topic prior to and during the meeting.

Kanouse et al. (1989) hold that "the panel should be well-informed about both the current state of science and the current state of practice." Panelists receive a bibliography on the consensus topic, including the paper abstracts from the bibliographic citations, at least one month in advance of the scheduled conference. The BID coordinator for each conference determines the scope and the nature of the literature search to be conducted through NLM. Background reports may also be prepared, and individual experts may be commissioned to compile summaries of the state of the science. According to Perry (1987), "there have been charges that data provided to some panels have been incomplete, that they have been overinterpreted or misrepresented, [and] that recommendations have been made in the absence of supporting evidence." At present, OMAR instructs panels that they may qualify the recommendations they make in the

consensus statement according to the strength of the evidence presented (Elliott, 1989).

Levitt et al. (1988) evaluated five papers used by panelists in the 1985 consensus development conference on adjuvant chemotherapy for breast cancer. The consensus statement favored the use of such treatment for breast cancer. The researchers found problems with the statistical analysis or in the performance of the randomized clinical trials in the five papers. These problems made them question the positive findings on the use of adjuvant chemotherapy. The researchers did not evaluate all of the articles used to formulate a consensus. Levitt et al. (1988) concluded:

> we feel that it is essential to ensure that a consensus of the use of these agents be based on careful evaluation not only of the conclusions, but also of the possible deficiencies of all the clinical trials aimed at establishing the values of a certain treatment.

Levitt et al. (1988) recommend that information retrieval for panelists should extend beyond a normal literature search, as the results of inconclusive or negative randomized clinical trials may not have been published.

Panelists may extrapolate from the evidence presented at the conference when they write the consensus statement. The 1984 consensus development conference on lowering blood cholesterol to prevent heart disease produced a contested consensus statement. Ahrens (1985) questioned the conference panel's recommendation of "a shift from the current typical American diet to one that is lower in total fat, saturated fat, and cholesterol" (NIH, 1985) for men, women, and children aged two and older. Ahrens questioned whether the available evidence warranted this recommendation and the safety and effectiveness of such a shift in diet for all Americans from age two.

Ahrens (1985) points out that "the panel leaned heavily on epidemiological evidence and public health considerations" as opposed to evidence from clinical trials. The association between high blood cholesterol levels and coronary heart disease is seen in cross-population comparison studies and in studies of migrants, although clinical trials of drugs and diets to lower blood cholesterol may not have established a significant difference in patient outcome (Kolata, 1985).

Both Ahrens (1985) and Kolata (1985) criticized the panel's use of the Coronary Primary Prevention Trial (CPPT). The CPPT was a randomized clinical trial comparing two groups of high-risk males, in which one group received cholestyramine and the other a placebo.

Ahrens and Kolata question the results of the trial. Comparisons of the experimental and control groups showed "a 19% reduction in risk ($P < .05$) of the primary endpoint-definite coronary heart disease (CHD) death and a 19% reduction in nonfatal myocardial infarction—reflecting a 24% reduction in definite CHD death and a 19% reduction in nonfatal myocardial infarction . . . . The risk of death from all causes was only slightly and not significantly reduced in the [experimental] cholestyramine group" (Lipid Research Clinics Program, 1984). The Lipids Research Clinic (LRC)-CPPT investigators note that "excess mortality in the LRC-CPPT cholestyramine group was confined to violent and accidental deaths. Since no plausible connection could be established between cholestyramine treatment and violent or accidental death, it is difficult to conclude that this could be anything but a chance occurrence" (Lipid Research Clinics Program, 1984).

Ahrens (1985) objects to the panel's "unjustifiable extrapolations" from clinical trials. Kolata (1985) and Ahrens (1985) questioned the panel's use of data on reduction of blood cholesterol in high-risk males, in whom little or no great health improvement was obtained, to support the consensus statement's broad recommendation for all Americans over age two to adhere to a diet "generally consistent with the most recent recommendations of the American Heart Association and the Atherosclerosis Study Group of the Inter-Society Commission on Heart Disease Resources" (NIH, 1985). In addition, Ahrens noted the absence of details on alternative diets for those who do attempt to lower blood cholesterol levels.

Oliver (1985) suggested that the CDP was highly imperfect for evaluation of the effect of cholesterol reduction on coronary heart disease. He noted specifically the shortage of time for the consensus process, the bias against cholesterol on the part of the panel, and the forced process of consensus in a highly controversial subject.

Oliver (1985) suggested alternative scenarios for technology assessment in order to better present and synthesize the available information:

- a learned body or government agency could establish a small group of experts to conduct an extended review of the literature and draft a report over the course of two years;
- experts could meet for a week of full discussion and a more complete literature review;

- a learned body or government agency could form a small standing committee to address topical issues in health care technology (Oliver, 1985).

These scenarios suggest that expansion of the time allocated for reaching consensus and for the preparation of the consensus statement would improve the outcome of consensus development activities (Oliver, 1985).

Development of Consensus

Some participants, including chairpersons, have been frustrated by the limited amount of time available for reaching consensus. Jacoby (1985) contended that "there is always the dilemma of not wanting to impose constraints (of time or rigidity) while at the same time trying to assure a useful outcome (i.e., closure) within an allotted time."

Fink et al. (1984) recommended that the level or type of consensus on any subject of debate should be defined in advance to ensure an efficient and agreeable settlement of disputes; for example,

- On the final vote, any topic supported by at least X% (a predetermined percentage) of participants is adopted.
- All topics are rated on a scale of 1 to 5, only those topics receiving a mean rating of 3.0 or greater are accepted.
- Any topic is dropped if it is vigorously opposed by at least X% (a predetermined percentage) of the participants (Fink et al., 1984).

Several consensus development conferences have experimented with decision models to help the panel explore the implications of the data presented by speakers (Jacoby and Pauker, 1986). McNeil and Pauker (1984) describe decision analysis as

> an exercise in building models—models that can be examined, modified, interrogated, and rebuilt—to provide insights into a choice in which outcomes are uncertain and risks are unavoidable. The underlying purpose of such an exercise would, on the surface, seem to be to arrive at the "correct" answer—the optimal strategy. But that is only a small part of the picture. The greater purpose is to provide a decision maker with an explicit reproducible process that can help structure complex alternatives in a rational way, a process that allows the incorporation of expertise and information from a variety of expert consultants without abdicating the decision to any one of them.

Jacoby and Pauker (1986) hold that the decision analytic process should involve five steps, which they describe as follows:

1. all of the available data are used to develop carefully structured, specific questions and to define the implications of the technology in terms of a logical sequence of potential strategies and outcomes;
2. the data are used to estimate the probability of each outcome;
3. the data are used to estimate the utility of each outcome;
4. consensus panelists use the model as an aid to determine the most desirable strategies; and
5. strategies are tested under different assumptions (sensitivity analysis).

Six consensus development conferences have made some use of this approach: diagnostic ultrasound in pregnancy (1984); postmenopausal estrogens in the prevention of osteoporosis (1984); limb-sparing treatment of adult soft-tissue and osteogenic sarcoma (1984); registries for bone marrow transplantation (1985); adjuvant chemotherapy for breast cancer (1985); and the impact of routine HTLV-III (human T-lymphotropic virus type III) antibody testing on public health (1986). Jacoby and Pauker (1986) presented Table A-1 to review the potential uses of decision analysis in consensus development. The panelists' use of the decision models has varied. Panelists followed the structured preposed model question areas for consideration to different degrees; some groups have relied upon the decision analysis in response to questions from the audience on the consensus statement (Jacoby and Pauker, 1986); other groups have made little use of the decision analysis. One NIH BID coordinator noted that the decision analysis did not add to the process of consensus development (Lipman, 1989). Jacoby and Pauker (1986) contend that the process should be refined and that planners should present information about the advantages and limitations of decision analysis in a more straightforward manner to better educate panelists.

The consensus process may not facilitate the expression of alternate viewpoints. Henig (1985) related the criticism that the "consensus statement—which almost by definition, leans toward the moderate, the mainstream, the prudent—can't embrace minority views." Several consensus conferences have produced the statements with

TABLE A.1 Decision Analysis in Consensus Development

Preconference	During Conference	Postconference
Planning	During presentations	Publication
Selecting topics	Modify model	Document reasoning
Formulating questions	Incorporate new data	Provide baseline for future assessments
Identifying needed input	Incorporate new assumptions	
Preparation	During deliberations	
Tuning panel in	Focus discussion	
Focusing speakers	Perform sensitivity analysis	
Providing consistent data	Integrate and interpret data	
	Make projections	
Evening before	During writing	
Identifying central issues	Focus consensus statement	
Broadening panel's views	Clarify specifics	
	During public discussion	
	Provide support against criticism	
	Permit justified modifications	

SOURCE: Israel Journal of Medical Sciences, 1986, v. 22, p. 186.

dissenting views (see, for example, the 1980 consensus development conference on cervical cancer screening). The majority of NIH consensus development conferences emerge with one statement accepted by the panels.

Preparation of Initial and Final Consensus Statements

Early in the program, critics were concerned that consensus statements would be "bland generalities that represent the lowest common denominator of a debate" (Rennie, 1981). Kahan et al. (1988) performed a content analysis of the statements of 24 consensus development conferences between 1979 and 1983. The researchers

recommended a didactic-style consensus statement, that is, one that is practical in orientation and offers detailed guidance for clinicians. The researchers hypothesize that concrete, differentiating recommendations (ones that suggest specific actions on the part of health care practitioners for different subclasses of patients) would influence physicians. Panels are currently asked to write a shorter section of conclusions and recommendations to emphasize the key points of the consensus statement.

There have been two attempts to involve a science writer in the composition of the consensus statement. In one instance, the science writer was brought in by the program planning committee; in the other, the chairperson selected the science writer to join the panel in executive session. Both times, the practice was reportedly unsuccessful as the panelists involved were unwilling to submit to outside editing and interference (Jacoby, 1988).

Dissemination of Consensus Information

According to OMAR, the dissemination of the results may now include the following:

A. Consensus development conferences usually receive considerable attention from the medical media and general media at the time of their occurrence. This serves to focus attention on the topic and the statement of the panel.
B. The OMAR Director of Communications and the BID Information Officer develop an information dissemination plan covering publicity for the conference and the strategy for distributing the Consensus Statement.
C. The Consensus Statement is printed by OMAR and distributed routinely to a variety of Federal health agencies, health care organizations, and the directors of continuing education of American Hospital Association membership hospitals. Additionally, the Consensus Statement is sent to targeted individuals and organizations specified in the information dissemination plan.
D. The *Journal of the American Medical Association* routinely publishes most of the Consensus Statements. Consensus Statements are also published by specialty journals in the area of the topic.
E. OMAR places notices in numerous professional journals announcing the availability of the Consensus Statement and inviting inquiry.

F. The publication of the Consensus Statement along with selected papers from a consensus development conference as a symposium is also a possibility. Proceedings of several conferences have been published in this manner either as supplements to specialty journals or as a monograph.

G. Summary videotapes and audiotapes of the conference may also be prepared and distributed.

H. A summary of the statement is also prepared and sent to appropriate specialty journals (OMAR, NIH, 1988).

References

Ahrens, E.H., Jr. 1985. The diet-heart question in 1985: Has it really been settled? The Lancet i:1085–1087.

Asch, S.M., and C.U. Lowe. 1984. The consensus development program: Theory, process, and critique. Knowledge: Creation, Diffusion, Utilization 5(3):369–385.

Elliott, J. 1989. Office of Medical Applications of Research, personal communication.

Fink, A., J. Kosecoff, M. Chassin, and R.H. Brook. 1984. Consensus methods: Characteristics and guidelines for use. The American Journal of Public Health 74(9):979–983.

Goodman, C., ed. 1988. Medical Technology Assessment Directory. Washington, D.C.: National Academy Press.

Henig, R.M. 1985. The process and provocations of consensus. Medical World News (October):96–111.

Institute of Medicine. 1985. Assessing Medical Technologies. Washington, D.C.: National Academy Press.

Jacoby, I. 1985. The consensus development program of the National Institutes of Health: Current practices and historical perspectives. International Journal of Technology Assessment in Health Care 1(2):420–432.

Jacoby, I. 1988. Uniformed Services University of Health Sciences, personal communication.

Jacoby, I., and S.G. Pauker. 1986. Technology assessment in health care: Group process and decision theory. Israel Journal of Medical Science 22:183–190.

Jacoby, I., and M. Rose. 1986. Transfer of information and its impact on medical practice: The U.S. experience. International Journal of Technology Assessment in Health Care 2(1):107–115.

Kahan, J.P., D.E. Kanouse, and J.D. Winkler. 1988. Stylistic variations in National Institutes of Health consensus statements, 1979–83. International Journal of Technology Assessment in Health Care 4(2):289–304.

Kanouse, D.E., R.H. Brook, J.D. Winkler, J. Kosecoff, S.H. Berry, G.M. Carter, J.P. Kahan, L. McCloskey, W.H. Rogers, C.M. Winslow, G.M. Anderson, L. Brodsley, A. Fink, and L. Meredith. 1989. Changing Medical Practice Through Technology Assessment: An Evaluation of the National Institutes of Health Consensus Development Program. Santa Monica, Calif.: The RAND Corporation.

Kolata, G. 1985. Heart panel's conclusions questioned. Science 227:40–41.

Levitt, S.H., R.A. Potish, D. Aeppli, and B. Lindgren. 1988. The consensus statements on adjuvant chemotherapy in breast cancer: Fact or artifact. American Journal of Oncology 11(1):73–76.

Lipids Research Clinics Program. 1984. The Lipid Research Clinics coronary primary prevention trial results. Journal of the American Medical Association 251:351–364.

Lipman, M. 1989. Georgetown University, Vincent Lombardi Cancer Center, personal communication.

Markle, G.E., and D.E. Chubin. 1987. Consensus development in biomedicine: The liver transplant controversy. The Milbank Quarterly 65(1):1–24.

McNeil, B.J., and S.G. Pauker. 1984. Decision analysis for public health: Principles and illustrations. Annual Review of Public Health 5:135–161.

Mullan, F., and I. Jacoby. 1985. The town meeting for technology. Journal of the American Medical Association 254(8):1068–1072.

National Institutes of Health. 1980. Cervical Cancer Screening: The Pap Smear. Consensus Development Conference Summary 3(4). Washington, D.C.: U.S. Government Printing Office.

National Institutes of Health. 1983. Charter of the NIH Coordinating Committee on Assessment and Transfer of Technology. Bethesda, Md.: National Institutes of Health. Photocopy.

National Institutes of Health. 1985. Lowering Blood Cholesterol to Prevent Heart Disease. Bethesda, Md.

Office of Medical Applications of Research, National Institutes of Health. 1988. Guidelines for the Selection and Management of Consensus Development Conferences . Bethesda, Md. Pamphlet.

Oliver, M.F. 1985. Consensus or nonsensus conferences on coronary heart disease. The Lancet i:1087–1089.

Perry, S. 1982. The brief life of the National Center for Health Care Technology. New England Journal of Medicine 307(17):1095–1100.

Perry, S. 1987. The NIH consensus development program: A decade later. New England Journal of Medicine 317:485–488.

Perry, S., and J.T. Kalberer, Jr. 1980. The NIH consensus-development program and the assessment of health-care technologies: The first two years. New England Journal of Medicine 303:169–172.

Rennie, D. 1981. Consensus statements. New England Journal of Medicine 304:665–666.

Wortman, P.M., and A. Vinokur. 1982. Evaluation of the National Institutes of Health Consensus Development Process, Phase I: Final Report. Ann Arbor: The Institute for Social Research, The University of Michigan.

Wortman, P.M., A. Vinokur, and L. Sechrest. 1988. Do consensus conferences work? A process evaluation of the National Institutes of Health consensus development program . Journal of Health Politics, Policy and Law 13(3):469–498.

B

NIH Consensus Development Conferences

Table B.1 provides a list of the National Institutes of Health (NIH) consensus development conferences.

TABLE B.1 NIH Consensus Development Conferences

Date	Conference Title	Cosponsor[a]
1977 September 14–16	Breast Cancer Screening	NCI
1978 May 22	Educational Needs of Physicians and Public Regarding Asbestos Exposure	NCI
June 13–14	Dental Implants: Benefit and Risk	NIDR
June 26–28	Mass Screening for Colo-Rectal Cancer	NCI
July 10–11	Treatable Brain Diseases in the Elderly	NIA
July 20	Indications for Tonsillectomy and Adenoidectomy: Phase I	NINCDS
September 14	Availability of Insect Sting Kits to Non-Physicians	NIAID
September 18–20	Mass Screening for Lung Cancer	NCI

Date	Conference Title	Cosponsor[a]
November 10-11	Supportive Therapy in Burn Care	NIGMS
December 4-5	Surgical Treatment of Morbid Obesity	NIAMDD[b]
1979		
February 16	Pain, Discomfort, and Humanitarian Care	(HHS)
March 5-7	Antenatal Diagnosis	NICHD
April 23-24	Transfusion Therapy in Pregnant Sickle Cell Disease Patients	NHLBI
April 26-27	Improving Clinical and Consumer Use of Blood Pressure Measuring Devices	NHLBI
June 5	The Treatment of Primary Breast Cancer: Management of Local Disease	NCI
June 27-29	Steroid Receptors in Breast Cancer	NCI
September 10-11	Intraocular Lens Implantation	NEI
September 13-14	Estrogen Use and Postmenopausal Women	NIA
October 15-16	Amantadine: Does it Have a Role in the Prevention and Treatment of Influenza?	NIAID
October 17-19	The Use of Microprocessor-Based "Intelligent" Machines in Patient Care	DRS
November 28-30	Removal of Third Molars	NIDR
1980		
April 10-12	Thrombolytic Therapy in Thrombosis	NHLBI
May 19-21	Febrile Seizures	NINCDS
July 14-16	Adjuvant Chemotherapy of Breast Cancer	NCI
July 23-25	Cervical Cancer Screening: The Pap Smear	NCI/NIA/NICHD
August 20-22	Endoscopy in Upper GI Bleeding	NIAMDD[b]
September 22-24	Cesarean Childbirth	NICHD
September 29-October 1	CEA as a Cancer Marker	NCI

Date	Conference Title	Cosponsor[a]
December 3-5	Coronary Artery Bypass Surgery: Scientific and Clinical Aspects	NHLBI
1981		
March 2-4	The Diagnosis and Treatment of Reye's Syndrome	NINCDS/NIAID/ NIDDK/ NICHD/ NIEHS/DRR
November 4-6	Computed Tomographic Scanning of the Brain	NINCDS/NCI
1982		
January 13-15	Defined Diets and Childhood Hyperactivity	NIAID/NICHD
March 1-3	Total Hip Joint Replacement	NIADDK[b]
November 1-3	Clinical Applications of Biomaterials	DRS
1983		
March 7-9	Critical Care Medicine	CC
June 20-23	Liver Transplantation	NIADDK[b]
September 27-29	Treatment of Hypertriglyceridemia	NHLBI
October 24-26	Precursors to Malignant Melanoma	NCI
November 15-17	Drugs and Insomnia: The Use of Medications to Promote Sleep	(NIMH)
December 5-7	Dental Sealants in the Prevention of Tooth Decay	NIDR
1984		
February 6-8	Diagnostic Ultrasound Imaging in Pregnancy	NICHD/DRR/(FDA)
February 27-29	Analgesic-Associated Kidney Disease	NIADDK[b]
April 2-4	Osteoporosis	NIADDK[b]
April 24-26	Mood Disorders: Pharmacologic Prevention of Recurrences	(NIMH)
September 24-26	Fresh Frozen Plasma: Indications and Risks	NHLBI/(FDA)

Date	Conference Title	Cosponsor[a]
December 3-5	Limb-Sparing Treatment of Adult Soft-Tissue and Osteosarcomas	NCI
December 10-12	Lowering Blood Cholesterol to Prevent Heart Disease	NHLBI
1985		
January 28-30	Travelers' Diarrhea	NIAID
February 11-13	Health Implications of Obesity	NIADDK[b]/NHLBI
April 22-24	Anesthesia and Sedation in the Dental Office	NIDR
June 10-12	Electroconvulsive Therapy	(NIMH)
September 9-11	Adjuvant Chemotherapy for Breast Cancer	NCI
1986		
January 13-15	Health Implications of Smokeless Tobacco Use	NCI/NIDR
March 24-26	Prevention of Venous Thrombosis and Pulmonary Embolism	NHLBI
May 19-21	Integrated Approach to the Management of Pain	CC
June 2-4	The Utility of Therapeutic Plasmapheresis for Neurological Disorders	NINCDS
July 7-9	Impact of Routine HTLV-III Antibody Testing of Blood and Plasma Donors on the Health of the Public	NHLBI
September 29-October 1	Infantile Apnea and Home Monitoring	NICHD
October 6-8	Platelet Transfusion Therapy	NHLBI
December 8-10	Diet and Exercise in Noninsulin-Dependent Diabetes Mellitus	NIDDK[b]
1987		
April 6-8	Newborn Screening for Sickle Cell Disease and Other Hemoglobinopathies	NHLBI
June 15-17	Management of Clinically Localized Prostate Cancer	NCI

Date	Conference Title	Cosponsor[a]
July 6-8	Differential Diagnosis of Dementing Diseases	NIA
July 13-15	Neurofibromatosis	NINCDS
October 19-21	Geriatric Assessment Methods for Clinical Decision Making	NIA
October 26-28	Magnetic Resonance Imaging	CC
1988		
March 28-30	Prevention and Treatment of Kidney Stones	NIDDK[b]
May 2-4	Cochlear Implants	NINCDS
June 13-15	Dental Implants	NIDR
June 27-29	Perioperative Red Cell Transfusion	NHLBI
October 3-5	Urinary Incontinence	NIA
1989		
March 6-8	Therapeutic Endoscopy and Bleeding Ulcers	NIDDK[b]
April 17-19	Oral Complications of Cancer Therapies: Diagnosis, Prevention, and Treatment	NIDR
May 8-10	Sunlight, Ultraviolet Radiation, and the Skin	NIAMS
September 13-15	Treatment of Destructive Behaviors in Persons with Developmental Disabilities	NICHD
1990		
January 22-24	Noise and Hearing Loss	NIDCD
March 19-21[c]	Surgery for Epilepsy	NINDS
March 26-28[c]	Treatment of Sleep Disorders in Older Persons	NIA/NHLBI/ NINDS/ (NIMH)
April/May[c]	Adjuvant Therapy for Colorectal Cancer	NCI
April/May[c]	Therapeutic Uses of Gammaglobulin	NIAID
June 18-21[c]	Treatment of Early-Stage Breast Cancer	NCI

Date	Conference Title	Cosponsor[a]
July[c]	Therapeutic Uses of Botulinum Toxin	NINDS
November[c]	Hyperparathyroidism	NIDDK[b]
December[c]	Treatment of Morbid Obesity	NIDDK[b]
1991		
January[c]	Melanoma	NCI

[a]Each conference is normally sponsored by OMAR and one or more bureau, institute, or division of NIH, which are abbreviated as follows.

CC	Warren Grant Magnuson Clinical Center
DRR	Division of Research Resources
DRS	Division of Research Services
NCI	National Cancer Institute
NEI	National Eye Institute
NHLBI	National Heart, Lung, and Blood Institute
NLM	National Library of Medicine
NIA	National Institute on Aging
NIADDK[b]	National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases (see note [b])
NIAID	National Institute of Allergy and Infectious Diseases
NIAMDD[b]	National Institute of Arthritis, Metabolism, and Digestive Diseases (see note [b])
NIAMS	National Institute of Arthritis and Musculoskeletal and Skin Diseases
NICHD	National Institute of Child Health and Human Development
NIDCD	National Institute on Deafness and Other Communication Disorders
NIDDK[b]	National Institute of Diabetes and Digestive and Kidney Diseases (see note [b])
NIDR	National Institute of Dental Research
NIEHS	National Institute of Environmental Health Sciences
NIGMS	National Institute of General Medical Sciences
NINCDS	National Institute of Neurological and Communicative Disorders and Stroke
NINDS	National Institute of Neurological Disorders and Stroke

Sponsors from organizations outside NIH are shown in parentheses and are abbreviated as follows:

FDA	Food and Drug Administration
HHS	U.S. Department of Health and Human Services
NIMH	National Institute of Mental Health

[b] Over time the name of the NIDDK has changed. From May 1972 to June 1981 the Institute was known as the NIAMDD, from June 1981 to April 1986 it was the NIADDK, and from April 1986 to the present it has been the NIDDK.

[c] These are the consensus conferences planned for 1990.

C

Questions from NIH Consensus Development Conferences (October 1987-September 1989)

Magnetic Resonance Imaging (MRI)

October 26, 1987

Are there contraindications to or risks of MRI?

What are the technological advantages and limitations (disadvantages) of MRI?

What are the clinical indications for MRI, and how does it compare to other diagnostic modalities?

What are the directions for future research in MRI?

Prevention and Treatment of Kidney Stones

March 30, 1988

What are the methods of medical prevention, and how successful are they?

What is the role of lithotripsy, and can it replace medical prevention?

What are the clinical and laboratory approaches for the evaluation of patients with stones?

What are the directions for future research?

Cochlear Implants

May 4, 1988

Who is a suitable candidate for a cochlear implant?

What are the advantages and disadvantages of the different types of cochlear implants?

How effective are cochlear implants?

What are the risks and limitations of cochlear implantation?

What are the special considerations for children?

What are the important directions for future research?

Dental Implants

June 15, 1988

What is the evidence that dental implants are effective for the long term?

What are the indications and contraindications of various types of dental implants?

What are the requirements for surgical, restorative, and periodontal management of patients with dental implants?

What are the health risks of dental implants?

What are the future directions for research on materials and designs of dental implants and on clinical management?

Perioperative Red Cell Transfusion

June 27–29, 1988

What should the criteria be for perioperative red blood cell transfusion?

What is the morbidity of anemia in the perioperative period?

What are the risks of red cell transfusion—both immediate and long term?

What are the alternatives to red cell transfusion?

What are the directions for future research?

Urinary Incontinence

October 3–5, 1988

What is the prevalence and clinical, psychological, and social impact of urinary incontinence among persons living at home and in institutions?

What are the pathophysiological and functional factors leading to urinary incontinence?

What diagnostic information should be obtained in assessment of the incontinent patient? What criteria should be employed to determine which tests are indicated for a particular patient?

What are the efficacies and limitations of behavioral, pharmacological, surgical, and other treatments for urinary incontinence

What sequences and/or combination of these interventions are appropriate? What management techniques are appropriate when treatment is not effective or indicated?

What strategies are effective in improving public and professional knowledge about urinary incontinence?

What are the needs for future research related to urinary incontinence?

Therapeutic Endoscopy and Bleeding Ulcers

March 6–8, 1989

Which patients with bleeding ulcers are at risk for rebleeding and thus emergency surgery?

How effective is endoscopic hemostatic therapy?

How safe is endoscopic hemostatic therapy?

Which bleeding patients should be treated?

What further research is required?

Oral Complications of Cancer Therapies: Diagnosis, Prevention, and Treatment

April 17–19, 1989

Is there a role for pretherapy interventions affecting the oral cavity in reducing the incidence of oral complications in the cancer patient?

Which pretreatment strategies are optimal to prevent or minimize oral complications?

What are the most effective strategies for management of acute oral complications occurring during cancer therapy?

What are the indicated strategies for management of chronic oral complications following cancer therapy?

What are the directions for future research?

Sunlight, Ultraviolet Radiation, and the Skin

May 8–10, 1989

What are the sources of ultraviolet radiation, and is the extent of human exposure changing over time?

What are the effects of sunlight on the skin?

What factors influence susceptibility to ultraviolet radiation?

Can ultraviolet-induced changes be prevented? If so, how?

Are sunlight-induced adverse skin alterations treatable and/or reversible? If so, how?

What are the directions for future research?

Treatment of Destructive Behaviors in Persons with Developmental Disabilities

September 11–13, 1989

What are the nature, extent, and consequences of destructive behaviors in persons with developmental disabilities?

What are the approaches to prevent, treat, and manage these behaviors?

What is the evidence that these approaches, alone or in combination, eliminate or reduce destructive behaviors?

What are the risks and benefits associated with the use of these approaches for the individual, family, and community?

Based on the answers to the above questions and taking into account (a) the behavior; (b) the diagnosis and functional level of the individual; (c) possible effects on the individual, family, and community; (d) the treatment setting; and (e) other factors, what recommendations can be made at present regarding the use of the different approaches?

What research is needed on approaches for preventing, treating, and managing destructive behaviors in persons with developmental disabilities?

D

General Bibliography on Consensus Development

Ahrens, E.H., Jr. 1985. The diet-heart question in 1985: Has it really been settled? The Lancet i:1085–1087.

Andreasen, P.B. 1988. Consensus conferences in different countries: Aims and perspectives. International Journal of Technology Assessment in Health Care 4:305–308.

Asch, S.M., and C.U. Lowe. 1984. The consensus development program: Theory, process, and critique. Knowledge: Creation, Diffusion, Utilization 5 (3):369–385.

Battista, R.N., and S.W. Fletcher. 1988. Making recommendations on preventative practices: Methodological issues. American Journal of Preventative Medicine 4(4):53–67.

Blumenthal, D. 1983. Federal policy toward health care technology: The case of the National Center. The Milbank Quarterly 61(4):584–613.

Bulpitt, C.J. 1988. Meta-analysis. The Lancet i:93–94.

Burke, J.F., H.S. Jordan, C.M. Boyle, and E. Vanner. 1981. An impact study of the 1978 consensus development conferences on supportive therapy in burn care. Washington, D.C.: U.S. Government Printing Office. Publication No. 0-134-132/3718.

Calltorp, J. 1988. Consensus development conferences in Sweden: Effects on health policy and administration. International Journal of Technology Assessment in Health Care 4:75–88.

Casparie, A.F., and J.E. van Everdingen. 1985. Consensus development conferences in The Netherlands. International Journal of Technology Assessment in Health Care 1:905–912.

Dalkey, N.C. 1969. The Delphi Method: An Experimental Study of Group Opinion. Santa Monica, Calif.: The RAND Corporation.

Delbecq, A., A.H. Van de Ven, and D.H. Gustafson. 1975. Group Techniques for Program Planning. Glenview, Ill.: Scott, Foresman.

Eddy, D.M. 1982. Clinical policies and the quality of clinical practice. New England Journal of Medicine 307:343–347.

Eddy, D.M. 1983. Flying without instruments: Analyzing medical policies by consensus and expert opinion. Bulletin. Winter. New York: The John A. Hartford Foundation.

Fink, A., J. Kosecoff, M. Chassin, and R.H. Brook. 1984. Consensus methods: Characteristics and guidelines for use. The American Journal of Public Health 74(9):979–983.

Gleicher, N. 1984. Cesarean section rates in the United States: The short-term failure of the national consensus development conference in 1980. Journal of the American Medical Association 252(23):3273–3276.

Goldbloom, R., R.N. Battista, and J. Haggerty. 1989. The periodic health examination. Canadian Medical Association Journal 141:205–208.

Goodman, C., ed. 1988. Medical Technology Assessment Directory. Washington, D.C.: National Academy Press.

Goodman, C., and S. Baratz, eds. 1990. Improving Consensus Development for Health Technology Assessment: An International Perspective. Washington, D.C.: National Academy Press.

Greer, A.L. 1987. The two cultures of biomedicine: Can there be a consensus? Journal of the American Medical Association 258(19)2739–2740.

Harris, S.S., C.J. Caspersen, G.H. Defriese, and E.H. Estes. 1989. U.S. preventive services task force. Journal of the American Medical Association 261 (24):3588–3589.

Henig, R.M. 1985. The process and provocations of consensus. Medical World News (October):96–111.

Hill, J.G. 1986. Technology assessment at the National Institute of Child Health and Human Development: An alternative model. Annals of the New York Academy of Sciences 477:351–355.

Hunter, J. 1982. Meta-analysis: Cumulating Research Findings Across Studies. Beverly Hills, Calif.: Sage Publications, Inc.

Institute of Medicine. 1985. Assessing Medical Technologies. Washington, D.C.: National Academy Press.

Jacoby, I. 1983. Biomedical technology information dissemination and the NIH consensus development process. Knowledge: Creation, Diffusion, Utilization 5(2):245–261.

Jacoby, I. 1985. The consensus development program of the National Institutes of Health: Current practices and historical perspectives. International Journal of Technology Assessment in Health Care 1(2):420–432.

Jacoby, I. 1988. Evidence and consensus. Journal of the American Medical Association 259(20):3039.

Jacoby, I. 1988. Update on assessment activities: United States perspective. International Journal of Technology Assessment in Health Care 4:95–105.

Jacoby, I., and S. Clark. 1986. Direct mailing as a means of disseminating NIH consensus statements. Journal of the American Medical Association 255 (10):1328–1330.

Jacoby, I., and S.G. Pauker. 1986. Technology assessment in health care: Group process and decision theory. Israel Journal of Medical Sciences 22:183–190.

Jacoby, I., and M. Rose. 1986. Transfer of information and its impact on medical practice: The U.S. experience. International Journal of Technology Assessment in Health Care 2(1):107–115.

Jennett, B. 1985. First consensus development program in the United Kingdom: On coronary artery bypass grafting , commentary by the chair of the conference. British Medical Journal 291:716–718.

Johnsson, M. 1988. Evaluation of the consensus conference program in Sweden, its impact on physicians. International Journal of Technology Assessment in Health Care 4:89–94.

Kahan, J.P., D.E. Kanouse, and J.D. Winkler. 1988. Stylistic variations in National Institutes of Health consensus statements, 1979–1983. International Journal of Technology Assessment in Health Care 4(2):289–304.

Kanouse, D.E., and I. Jacoby. 1988. When does information change practitioners' behavior? International Journal of Technology Assessment in Health Care 4 (1):27–33.

Kanouse, D.E., R.H. Brook, J.D. Winkler, J. Kosecoff, S.H. Berry, G.M. Carter, J.P. Kahan, L. McCloskey, W.H. Rogers, C.M. Winslow, G.M. Anderson, L. Brodsley, A. Fink, and L. Meredith. 1989. Changing Medical Practice Through Technology Assessment: An Evaluation of the National Institutes of Health Consen

sus Development Program. Santa Monica, Calif.: The RAND Corporation.

Klazinga, N.S., A.F. Casparie, and J.E. van Everdingen. 1987. Contribution of medical decision-making to consensus development conferences. Health Policy 8:339–346.

Kolata, G. 1985. Heart panel's conclusions questioned. Science 227:40–41.

Kosecoff, J., D.E. Kanouse, W.H. Rogers, L. McCloskey, C.M. Winslow, and R.H. Brook. 1987. Effects of the National Institutes of Health consensus development program on physician practice. Journal of the American Medical Association 258(19):2708–2713.

Levitt, S.H., R.A. Potish, D. Aeppli, and B. Lindgren. 1988. The consensus statements on adjuvant chemotherapy in breast cancer: Fact or artifact. American Journal of Oncology 11(1):73–76.

Linstone, H., and M. Turoff, eds. 1975. The Delphi Method: Techniques and Applications. Reading, Mass: Addison-Wesley.

Lipids Research Clinics Program. 1984. The Lipid Research Clinics coronary primary prevention trial results. Journal of the American Medical Association 251:351–374.

Lomas, J. 1986. The consensus process and evidence dissemination. Canadian Medical Association Journal 134:1340–1341.

Lomas, J., G. Anderson, M. Enkin, E. Vayda, R. Roberts, and B. MacKinnon. 1988. The role of evidence in the consensus development process. Knowledge: Creation, Diffusion, Utilization 5(2):245–261.

Louis, T.A., H.V. Fineberg, and F. Mosteller. 1985. Findings for public health from meta-analyses. Annual Review of Public Health 6:1–20.

Maier, N. 1963. Leadership principles for problem solving conferences. Modification of pp. 239–253 in Maier, N., Problem Solving Discussions and Conferences, New York: McGraw-Hill (first appeared in Michigan Business Review, pp. 8–15, May 1982).

Maier, N. 1967. Assets and liabilities in group problem solving: The need for an integrative function. Psychological Review 74:239–249.

Markle, G.E., and D.E. Chubin. 1987. Consensus development in biomedicine: The liver transplant controversy. The Milbank Quarterly 65(1):1–24.

McGregor, M. 1989. Technology and the allocation of resources. New England Journal of Medicine 320(2):118–120.

McNeil, B.J., and S.G. Pauker. 1984. Decision analysis for public health: Principles and illustrations. Annual Review of Public Health 5:135–161.

Mullan, F., and I. Jacoby. 1985. The town meeting for technology. Journal of the American Medical Association 254(8):1068–1072.

National Institute of Child Health and Human Development. 1989. Pp. 1–10 of Chapter 1 of draft report on treatment of destructive behaviors in persons with developmental disabilities. Bethesda, Md. Photocopy.

National Institutes of Health. 1980. Cervical Cancer Screening: The Pap Smear. Consensus Development Conference Summary 3(4). Washington, D.C.: U.S. Government Printing Office.

National Institutes of Health. 1983. Charter of the NIH Coordinating Committee on Assessment and Transfer of Technology. Bethesda, Md.: National Institutes of Health. Photocopy.

National Institutes of Health. 1984. Dental Sealants in the Prevention of Tooth Decay. Consensus Development Conference Summary 4(11). Washington, D.C.: U.S. Government Printing Office.

National Institutes of Health. 1985. Lowering Blood Cholesterol to Prevent Heart Disease. Bethesda, Md.

Office of Medical Applications of Research, National Institutes of Health. Undated. Office of Medical Applications of Research. Bethesda, Md. Pamphlet.

Office of Medical Applications of Research, National Institutes of Health. 1988. Guidelines for the Selection and Management of Consensus Development Conferences. Bethesda, Md. Pamphlet.

Oliver, M.F. 1985. Consensus or nonsensus conferences on coronary heart disease. The Lancet i:1087–1089.

Olsen, S.A. 1982. Group Planning and Problem Solving Methods in Engineering Management. New York: John Wiley & Sons.

Pauker, S. 1986. Decision analysis as a synthetic tool for achieving consensus in technological assessment. International Journal of Technology Assessment in Health Care 2(1):83–97.

Perry, S. 1982. The brief life of the National Center for Health Care Technology. New England Journal of Medicine 307(17):1095–1100.

Perry, S. 1987. The NIH consensus development program: A decade later. New England Journal of Medicine 317:485–488.

Perry, S. 1988. Consensus development: An historical note. International Journal of Technology Assessment in Health Care 4:481–488.

Perry, S., and J.T. Kalberer, Jr. 1980. The NIH consensus-development program and the assessment of health-care technologies: The first two years. New England Journal of Medicine 303:169–172.

Pillemer, D.B. and R.J. Light. 1980. Synthesizing outcomes: How to use research evidence from many sources. Harvard Educational Review 50(2):176–195.

Policy Research Incorporated. 1977. A Comprehensive Study of the Ethical, Legal, and Social Implications of Advances in Biomedical and Behavioral Research and Technology: Summary of the Final Report. Baltimore.

Porter, A.L., F.A. Rossini, S.R. Carpenter, and A.T. Roper. 1980. A Guidebook for Technology Assessment and Impact Analysis. New York: North-Holland.

Public Health Service. 1989. Caring for Our Future: The Content of Prenatal Care. Report of the Public Health Service Expert Panel on the Content of Prenatal Care. Washington, D.C: U.S. Department of Health and Human Services.

Rennie, D. 1981. Consensus statements. New England Journal of Medicine 304:665–666.

Riesenberg, D. 1987. Consensus conferences. Journal of the American Medical Association 258(19):2738.

Sackett, D.L. 1989. Rules of evidence and clinical recommendations on the use of antithrombotic agents. Chest 95:2S–4S.

Sacks, H.S., J. Berrier, D. Reitman, V.A. Ancona-Berk, and T.C. Chalmers. 1987. Meta-analyses of randomized controlled trials. New England Journal of Medicine 316(8):450–455.

Stocking, B. 1985. First consensus development conference in United Kingdom: On coronary artery bypass grafting, views of audience, panel, and speakers. British Medical Journal 291:713–716.

Thacker, S.B. 1988. Meta-analysis: A quantitative approach to research integration. Journal of the American Medical Association 259(11):1685–1689.

Vang, J. 1986. The consensus development conference and the European experience. International Journal of Technology Assessment in Health Care 2:65–76.

Winkler, J.D., D.E. Kanouse, L. Brodsley, and R.H. Brook. 1986. Popular press coverage of eight National Institutes of Health consensus development topics. Journal of the American Medical Association 255(10):1323–1327.

Wortman, P.M., and A. Vinokur. 1982. Evaluation of the National Institutes of Health Consensus Development Process, Phase I: Final Report. Ann Arbor: The Institute for Social Research, The University of Michigan.

Wortman, P.M., A. Vinokur, and L. Sechrest. 1988. Do consensus conferences work? A process evaluation of the NIH consensus development program. Journal of Health Politics, Policy and Law 13(3):469–498.